



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 99448**

**TO: Karen A Lacourciere**  
**Location: CM1/11D09/11E12**  
**Art Unit : 1635**  
**Friday, July 25, 2003**

**Case Serial Number: 08765244**

**From : Susan Hanley**  
**Location: Biotech-Chem Library**  
**CM1 6B05**  
**Phone: 305-4053**

**[susan.hanley@uspto.gov](mailto:susan.hanley@uspto.gov)**

### **Search Notes**



# STIC SEARCH RESULTS FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor  
308-4258, CM1-1E01

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 – Circ. Desk



STIC-Biotech/ChemLib

99448

RECEIVED

JUL 23 2003

(STIC)

**From:** Lacourciere, Karen  
**Sent:** Tuesday, July 22, 2003 5:29 PM  
**To:** STIC-Biotech/ChemLib  
**Subject:** sequence search request 08/765,244

Please search SEQ ID NO:1 and 22 in the amino acid databases for 08/765,244. Thank-you!

*Karen A. Lacourciere Ph.D.*

CM1 11D09 GAU 1635  
(703) 308-7523

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 21:59:18 ; Search time 36.119 seconds  
(without alignments)  
180.176 Million cell updates/sec

Title: US-08-765-244-1  
Perfect score: 208  
Sequence: 1 MSLNRLILLKAALKRAHTS.....VRNFRYKPVQSQLKPRDLC 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :	A_Geneset_19Jun03.*	
	1: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*	
	2: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*	
	3: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*	
	4: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*	
	5: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*	
	6: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*	
	7: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*	
	8: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*	
	9: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*	
	10: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*	
	11: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*	
	12: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*	
	13: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*	
	14: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*	
	15: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*	
	16: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*	
	17: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*	
	18: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*	
	19: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*	
	20: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*	
	21: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*	
	22: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*	
	23: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*	
	24: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	197	94.7	43	17 AAR90584	Rat ornithine tran
2	159	76.4	32	23 ABG30857	Rat ornithine tran
3	108	51.9	32	23 ABG30856	Human ornithine tr
4	108	51.9	258	21 AAB18445	Protein encoded by
5	105	50.5	32	22 AAG64224	OTC peptide fragme
6	92	44.2	32	15 AAR48260	Mitochondrial matr
7	92	44.2	32	20 AAW96358	Mitochondrial matr
8	92	44.2	32	23 ABG92993	Localisation sequ
9	92	44.2	32	24 ABP56588	Mitochondrial matr

10	79.5	38.2	31	21 AAB15704	Mitochondrial matr
11	79.5	38.2	31	21 AAB22835	Mitochondrial matr
12	58	27.9	165	20 AAY04933	Mycobacterium spec
13	55.5	26.7	149	22 AAU23468	Novel human enzyme
14	55.5	26.7	761	22 AAB46718	S. cerevisiae DNA
15	54.5	26.2	138	21 AAY75465	Neisseria meningit
16	54.5	26.2	138	21 AAY75466	Neisseria meningit
17	51.5	24.8	280	21 AAG04153	Arabidopsis thalia
18	51.5	24.8	283	21 AAG04152	Arabidopsis thalia
19	51.5	24.8	303	21 AAG26195	Arabidopsis thalia
20	51.5	24.8	303	21 AAG37586	Arabidopsis thalia
21	51.5	24.8	306	21 AAG26194	Arabidopsis thalia
22	51.5	24.8	306	21 AAG37585	Arabidopsis thalia
23	51	24.5	135	6 AAP30637	Salmon pleopionela
24	50.5	24.3	296	22 ABG64170	Drosophila melanog
25	50.5	24.3	755	22 ABG59259	Drosophila melanog
26	50	24.0	136	22 AAB74673	Human protease and
27	50	24.0	193	23 ABB89739	Human polypeptide
28	50	24.0	250	23 ABB99949	Dipeptidyl peptida
29	50	24.0	310	22 AAB47188	Human DPP8 524phe-
30	50	24.0	310	23 ABB08994	Human dipeptidyl p
31	50	24.0	465	22 AAB47189	Human DPP8 319thr-
32	50	24.0	593	20 AAY50124	Human vesicle tran
33	50	24.0	593	20 AAW70702	Human vesicle tran
34	50	24.0	593	22 AAB31567	Amino acid sequenc
35	50	24.0	632	22 AAB93565	Human protein sequ
36	50	24.0	724	23 ABB97362	Novel human protei
37	50	24.0	782	23 ABB97361	Novel human protei
38	50	24.0	866	22 ABB57741	Drosophila melanog
39	50	24.0	882	22 AAB47187	Human DPP8. Homo
40	50	24.0	882	23 AAB24170	Human dipeptidyl p
41	50	24.0	882	23 ABB61591	Human DPPIV relate
42	50	24.0	882	23 AAW74749	Human protease PRT
43	50	24.0	882	23 AAG78415	Amino acid sequenc
44	50	24.0	882	24 ABO07720	Human serine prote
45	49.5	23.8	115	23 ABP02239	Human ORFX protein

ALIGNMENTS

RESULT 1	
AAR90584	
ID	AAR90584 standard; Protein; 43 AA.
AC	AAR90584;
DT	25-MAR-2003 (updated)
DT	31-OCT-1996 (first entry)
DE	Rat ornithine transcarbamylase signal peptide.
KW	promoter; peptide-nucleic acid; cyclised; gene therapy; target;
KW	site-directed mutagenesis; introduction; protein transport.
OS	Synthetic.
PN	DE19520815-A1.
PD	21-DEC-1995.
PF	11-JUN-1995; 95DE-1020815.
PR	16-JUN-1994; 94DE-4421079.
PA	(SEIB/) SEIBEL P.
PI	Seibel A, Seibel P;
DR	WPI; 1996-041226/05.
PT	Replicable and transcriptionally active plasmid carrying signal
PT	peptide for specific target - useful for site directed mutagenesis



XX The invention relates to a recombinant vector for expressing a fused  
CC protein containing a fused gene in which a base sequence defining a  
CC transfer signal peptide to small cellular organs having an extranuclear  
CC gene is combined with a base sequence defining the amino acid sequence of  
CC a restriction enzyme recognising a defined base sequence. The vector is  
CC used for decomposing an extranuclear gene of a nonhuman organism.  
CC The present sequence represents the signal peptide of rat mitochondrial  
CC ornithine transcarbamylase (MOT) which may be used in the vector of the  
CC invention.

07-OCT-2002	(first entry)
Human ornithine transcarbamylase	signal peptide.
Human; signal peptide;	ornithine transcarbamylase; MOF;
recombinant vector;	fusion protein; extranuclear gene.

25-JUN-2002.

14-DEC-2000; 2000JP-0380975.

14-DEC-2000; 2000JP-0380975.

(TANA/) TANAKA M.  
(GIFU-) ZH GIFUKEN KOKUSAI BIO KENYUJO.  
(OYOS-) ZH OYO SEIRAGAKU KENYUSHO.

WPI: 2002-569946/61.  
N-PSDB; ABK88418.

A recombinant vector for expressing a fused protein, useful for decomposing an extranuclear gene of a nonhuman organism -

Disclosure; Page 3; 15pp; Japanese.

The invention relates to a recombinant vector for expressing a fused protein containing a fused gene in which a base sequence defining a transfer signal peptide to small cellular organs having an extranuclear gene is combined with a base sequence defining the amino acid sequence of a restriction enzyme recognising a defined base sequence. The vector is used for decomposing an extranuclear gene of a nonhuman organism. The present sequence represents the signal peptide of human mitochondrial ornithine transcarbamylase (MOT) which may be used in the vector of the invention.

Query Match 51.9%; Score 108; DB 23; Length 32;  
Best Local Similarity 68.8%; Pred. No. 1.1e-09;  
Matches 22; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

The specification describes a method for functionally complementing one or more defects, mutations, or deletions in a mitochondrial genome of a cell having a nuclear genome. The method is used for treating a disease or disorder that arises from deletion of the protein-encoding genes of the mitochondrial genome. Alternatively, the method is used for treating a disease or disorder that arises from one or more defects, deletions or mutations in mitochondrial genes encoding ribosomes or tRNA for transcription and translation in the mitochondria. The diseases or disorders that can be ameliorated are mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes, Leber hereditary optic neuropathy, myoclonic epilepsy with ragged-red fibres, neurogenic muscular weakness, ataxia, retinitis pigmentosa, Kearns-Sayre syndrome; Leigh syndrome, Pearson Marrow pancreas syndrome, amnioglycoside-associated deafness, diabetes with deafness, leukodystrophy with hypotonia, autism with seizures, sudden infant death syndrome with hypoglycemia, leukaemia with maternally inherited thrombocytopenia, migraines (associated with hearing loss, strokes, or diabetes), early hearing loss, refractory infantile reflux with carnitine deficiency, multiple sclerosis with seizures, blindness with optic atrophy and dystonia, renal tubular acidosis with elevated lactic acid and hypotonia, nonvalvular hypertrophic cardiomyopathy before age 50 and chronic pancreatitis with stroke-like episodes. The present sequence is encoded by plasmid pUOATP2. Plasmid pUOATP2 comprises a mutant oligomycin-

```

XX DE Mitochondrial matrix retention signal.
XX DE
XX KW Single chain antibody; sFv; heavy chain; light chain;
XX KW variable domain; hydrophilic linker; antibodies; targeting;
XX KW subcellular localisation signal; mitochondrial matrix;
XX KW retention signal.
XX OS Synthetic.
XX PH Key Location/Qualifiers
XX FT Misc-difference 7 /note= "not defined"
XX FT Misc-difference 8 /note= "not defined"
XX FT Misc-difference 32 /note= "not defined"
XX PN WO9402610-A1.
XX PD 03-FEB-1994.
XX PF 16-JUL-1993; 93WO-US06735.
XX PR 17-JUL-1992; 92US-0916939.
XX PR 17-MAR-1993; 93US-0045274.
XX PA (DAND ) DANA FARBER CANCER INST INC.
XX PI Haseltine WA, Marasco WA;
XX DR WPI; 1994-048868/06.
XX PT Intracellular binding of antigens - by using antibody targeting
XX PT with vector system, for e.g. tumour suppression
XX PS Disclosure; Page 103; 155pp; English.
XX CC New vector systems comprise a sequence adapted for intracellular
XX CC delivery and expression contg. a promoter operably linked to an
XX CC antibody gene encoding an antibody which binds to a specific target
XX CC antigen. The antibody is esp. a single chain antibody in which the
XX CC heavy and light chain variable regions are joined via a hydrophilic
XX CC linker peptide. Localisation sequences are pref. included in the
XX CC constructs. The sequence AAR48260 is a mitochondrial matrix retention
XX CC signal.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 32 AA;
Query Match 44.2%; Score 92; DB 15; Length 32;
Best Local Similarity 61.3%; Pred. No. 3.9e-07;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
QY 1 MLSNLRILLNKAALRKHAHTSMVRNFRYKGPV 31
DB 1 MLFNLRXXLLNNAAFRRGHNFVRNFRCGQPL 31
RESULT 7
ID AAW96358
XX AAW96358 standard; Peptide; 32 AA.
XX AC AAW96358;
XX DT 19-JUL-1999 (first entry)
XX DE Mitochondrial matrix localisation signal peptide.
XX KW Antibody; immune response; modulation; MHC; IRM; receptor;
XX KW intrabody; major histocompatibility complex; graft rejection;
XX KW immunomodulatory response molecule; regulation; transplantation;
XX KW retention signal; localisation signal; golgi apparatus; ER;

```

```

KW endoplasmic reticulum.
XX OS Synthetic.
XX PH Key Location/Qualifiers
XX FT Misc-difference 7 /note= "Any amino acid"
XX FT Misc-difference 8 /note= "Any amino acid"
XX FT Misc-difference 32 /note= "Any amino acid"
XX PN WO9914353-A2.
XX PD 25-MAR-1999.
XX PF 18-SEP-1998; 98WO-US19563.
XX PR 19-SEP-1997; 97US-0059339.
XX PA (DAND ) DANA FARBER CANCER INST INC.
XX PI Marasco W, Mhashikar A;
XX DR WPI; 1999-229546/19.
XX PT Altering the regulation of the immune system
XX PS Disclosure; Page 28; 56pp; English.
XX CC Intracellular binding to a desired target by an intracellularly
XX CC expressed antibody (i.e. an intrabody) can be used to knock out
XX CC multiple locuses of immunomodulatory receptor molecules (IRMs),
XX CC so that the expression of multiple major histocompatibility (MHC)
XX CC molecules is blocked. This selective targeting of IRMs, their
XX CC pathways or components, can be used to selectively regulate the
XX CC immune system by controlling expression of these molecules and
XX CC preventing an undesired immune response in a cell. Any component
XX CC of the MHC pathway or the MHC assembly line or antigen presentation
XX CC can be targeted. Intrabodies can be used to knock out the immune
XX CC response in a particular tissue or portion of the body to prepare
XX CC it for cell or tissue transplantation. Alternatively, an organ for
XX CC transplantation can be perfused with the intrabody ex vivo. The
XX CC intrabodies can comprise whole antibodies, heavy chains, Fab'
XX CC fragments, single-chain antibodies and diabodies. The intrabodies
XX CC also comprise an intracellular localisation signal to facilitate
XX CC interception of expressed proteins. For example, if the target was
XX CC a cell surface receptor, the antibody would comprise a leader
XX CC sequence and an endoplasmic reticulum (ER) or Golgi apparatus
XX CC retention signal. This peptide is a localisation sequence for the
XX CC mitochondrial matrix. For other localisation sequences see
XX CC AAW96345-W96377.
XX SQ Sequence 32 AA;
Query Match 44.2%; Score 92; DB 20; Length 32;
Best Local Similarity 61.3%; Pred. No. 3.9e-07;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
QY 1 MLSNLRILLNKAALRKHAHTSMVRNFRYKGPV 31
DB 1 MLFNLRXXLLNNAAFRRGHNFVRNFRCGQPL 31
RESULT 8
ID AAW92993
XX AAW92993 standard; Peptide; 32 AA.
XX AC AAW92993;
XX DT 20-NOV-2002 (first entry)
XX DE Localisation sequence to direct antibodies to the mitochondria.

```

Regulator; transcription; cell death; phenotype; molecular scaffold; gene therapy; cancer; cardiovascular disease; arrhythmia; heart failure; ischemia; obesity; neurodegenerative disease; Alzheimer's disease; bone pathology; dermatologic disease; psoriasis; infection; AIDS; acquired immunodeficiency syndrome; cosmetic; wound healing; antibiotic transport; drug toxicity; drug resistance; immunobiology; inflammation; allergic response; human immunodeficiency virus.

Unidentified.

WO200262822-A2.

15-AUG-2002.

04-FEB-2002; 2002WO-US02914.

02-FEB-2001; 2001US-265589P.

05-FEB-2001; 2001US-265880P.

27-FEB-2001; 2001US-371423P.

23-JAN-2001; 2001US-263226P.

28-MAR-2000; 2000US-192586P.

22-SEP-1997; 97US-935377P.

(UYRP ) UNIV ROCHESTER.

Zauderer M, Smith ES;

WPI; 2002-643398/69.

Identifying regulator polypeptides which influence target transcriptional regulatory regions, useful for treating cancer, comprises introducing host cells expressing the polypeptide into a library of polynucleotides -

Disclosure; Page 37; 224pp; English.

The invention discloses a method for identifying polynucleotides encoding a regulator polypeptide, whose expression induces activation of a target transcriptional regulatory region in a host cell. The method comprises providing a population of eukaryotic host cells capable of expressing the polypeptide, introducing into the host cell a library of polynucleotides encoding the polypeptides, permitting expression of the polypeptides and then recovering them from the host cells. The target transcriptional regulatory region is operably associated with a polynucleotide encoding a gene product, the expression of which results in host cell death or cause the host cells to exhibit a pre-determined modified phenotype and where the gene product is expressed upon activation of target transcriptional regulatory region. Each candidate regulator polypeptide comprises a candidate peptide and a molecular scaffold fused to the peptide so that the peptide is displayed on the surface of the candidate regulator polypeptide. The methods are useful in selecting and/or screening regulator molecules, such as polypeptides, which directly or indirectly induce or suppress the transcriptional activation of a target transcriptional regulatory region in a eukaryotic host cell. These regulator molecules may be used (e.g. in gene therapy) for preventing or treating cancers (e.g. breast or ovarian cancer), cardiovascular diseases (e.g. arrhythmia, heart failure, ischemia), obesity, neurodegenerative diseases (e.g. Alzheimer's disease), bone pathologies, dermatologic diseases (e.g. psoriasis), infections (e.g. viral, bacterial), acquired immunodeficiency syndrome (AIDS), in cosmetic applications and in wound healing. The method is also useful in screening regulator molecules that block antibiotic transport mechanisms, in drug toxicities and drug resistance applications and in improving the performance of existing or developmental drugs. It may also be used in immunobiology, inflammation, allergic response and in biotechnology applications. The sequences presented in ABG92946-ABG93029 are examples of regulator polypeptides.

Sequence 32 AA;

Query Watch 44.2%; Score 92; DB 23; Length 32;

Best Local Similarity 61.3%; Pred. No. 3.9e-07;

Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MSLNRLILLNKAALFKAHTSMVRNFRYKGPV 31

DB 1 MLFNLRXXLNNAAPRHGHNFVNRFCGQPL 31

RESULT 9

ABP56588

ID ABP56588 standard; Peptide; 32 AA.

XX

AC ABP56588;

XX

DT 24-WAR-2003 (first entry)

XX

DE Mitochondrial matrix targeting peptide SEQ ID NO:54.

XX

KW Identification; intrabody; eukaryotic cell; immunoglobulin; selection;

KW cardiovascular; diminished arrhythmia potential; cardiomyocyte; stroke;

KW enhanced contractile property; heart failure; arrhythmia; embolic;

KW sarcolemmal calcium cycling; artery; arteriole; angina; atherosclerosis;

KW LDL metabolism; HDL metabolism; skin biology; keloid formation.

XX

OS Unidentified.

XX

FH Key Location/Qualifiers

FT Misc-difference 7 /note= "any amino acid"

FT FT

FT Misc-difference 8 /note= "any amino acid"

FT FT

FT Misc-difference 32 /note= "any amino acid"

FT FT

PN WO200286096-A2.

XX

PD 31-OCT-2002.

XX

PF 23-JAN-2002; 2002WO-US01677.

XX

PR 23-JAN-2001; 2001US-263225P.

PR 24-JAN-2001; 2001US-263220P.

PR 27-FEB-2001; 2001US-271422P.

PR 15-JUN-2001; 2001US-298095P.

XX

PA (UYRP ) UNIV ROCHESTER MEDICAL CENT.

XX

PI Zauderer M, Wei C, Smith ES;

XX

DR WPI; 2003-103408/09.

XX

PT Selecting polynucleotides encoding an intracellular immunoglobulin

PT which induces a modified phenotype in a eukaryotic host cell, by

PT introducing library of polynucleotides encoding immunoglobulin subunit

PT polypeptides -

PS Disclosure; Page 44; 257pp; English.

XX

CC The present invention describes a method for selecting polynucleotides

CC (PNS) encoding an intracellular immunoglobulin molecule or its fragment

CC whose expression induces a modified phenotype in a eukaryotic host cell

CC (I). The method comprises introducing into (I) a first and second library

CC of PNS encoding, through operable association with a transcriptional

CC control region, first and second intracellular immunoglobulin subunit

CC polypeptides, respectively. The method is useful for selecting

CC polynucleotides which encode an intracellular immunoglobulin molecule, or

CC fragment. The method is useful e.g. for identifying polynucleotides which

CC singly or collectively encode intracellular immunoglobulin molecules, or

CC which sensitize host cells to killing by an agent. The method may also be

CC used in cardiovascular applications; for screening for diminished

CC arrhythmia potential in cardiomyocytes and for enhanced contractile

CC properties of cardiomyocytes and diminish heart failure potential; for

CC identifying intracellular immunoglobulin molecules that will regulate

CC intracellular and sarcolemmal calcium cycling in cardiomyocytes to

CC prevent arrhythmias or that will diminish embolic phenomena in arteries



diabetes, autoimmune diseases, inflammatory diseases, infectious diseases and especially cancer. The delivery method of the invention causes a drug to become almost immediately localised at its site of action in an inactive form where it accumulates. Once released by antibody degradation, the drug is at an effective concentration only at the target site, with very little free drug being available in the rest of the body. The system of the invention therefore reduces the side effects caused by therapeutic agents, and also provides economic benefits as a smaller amount can be administered to the patient. Sequences AAB22833- AAB22837 and AAB22839-822857 represent subcellular localisation sequences which can be incorporated into a drug-specific antibody used in the method of the invention.

Query Match 38.2%; Score 79.5; DB 21; Length 31;  
Best Local Similarity 61.3%; Pred. No. 3.6e-05;  
Matches 19; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

**Qy**     1 MLNLRIILLNKAALRKAHTSMVRNFYKPV 31  
          || ||| || | | | | | | | | | :  
**Dd**     1 MLFNLR-XLNNAAFRHGHNFMRNFCGGPL 30

Query Match 27.9%; Score 58; DB 20; Length 165;  
Best Local Similarity 39.4%; Pred. No. 0.73;

Matches	13;	Conservative	7;	Mismatches	13;	Indels	0;	Gaps	0;
Qy	6	RIILNKAAALRKHAHTSMVRNFRYKPVQSQLKPR	38						
		:      :        :							
Db	119	RVILRCATHRKANQSRARTLRPLRLPLRALRPR	151						
RESULT_13									
AAU23468									
ID	AAU23468 standard; Protein: 149 AA.								
XX	AC								
XX	AAU23468;								
XX	18-DEC-2001	(first entry)							
XX	Novel human enzyme polypeptide #554.								
DE									
XX									
KW	Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;								
KW	ligase; hyperproliferative disorder; immunodeficiency disorder;								
KW	autoimmune disorder; neurological disorder; metabolic disorder;								
KW	inflammatory disorder; cardiovascular disorder; reproductive disorder;								
KW	blood-related disorder; infectious disorder; cytostatic; anti arthritic;								
KW	nephrotropic; anticoagulant.								
XX									
OS	Homo sapiens.								
XX									
PN	WO200155301-A2.								
XX									
PD	02-AUG-2001.								
XX									
PF	17-JAN-2001; 2001WO-US01239.								
XX									
PR	31-JAN-2000; 2000US-0179065.								
PR	04-FEB-2000; 2000US-0180828.								
PR	24-FEB-2000; 2000US-0184664.								
PR	02-MAR-2000; 2000US-0186350.								
PR	16-MAR-2000; 2000US-0189874.								
PR	17-MAR-2000; 2000US-0190076.								
PR	18-APR-2000; 2000US-0198123.								
PR	19-MAY-2000; 2000US-0205515.								
PR	07-JUN-2000; 2000US-0209467.								
PR	28-JUN-2000; 2000US-0214886.								
PR	30-JUN-2000; 2000US-0215135.								
PR	07-JUL-2000; 2000US-0216647.								
PR	07-JUL-2000; 2000US-0216880.								
PR	11-JUL-2000; 2000US-0217487.								
PR	11-JUL-2000; 2000US-0217496.								
PR	14-JUL-2000; 2000US-0218290.								
PR	26-JUL-2000; 2000US-0220963.								
PR	28-JUL-2000; 2000US-0220964.								
PR	14-AUG-2000; 2000US-0224518.								
PR	14-AUG-2000; 2000US-0224519.								
PR	14-AUG-2000; 2000US-0225213.								
PR	14-AUG-2000; 2000US-0225214.								
PR	14-AUG-2000; 2000US-0225266.								
PR	14-AUG-2000; 2000US-0225267.								
PR	14-AUG-2000; 2000US-0225268.								
PR	14-AUG-2000; 2000US-0225270.								
PR	14-AUG-2000; 2000US-0225447.								
PR	14-AUG-2000; 2000US-0225757.								
PR	14-AUG-2000; 2000US-0225758.								
PR	14-AUG-2000; 2000US-0225759.								
PR	18-AUG-2000; 2000US-0226279.								
PR	22-AUG-2000; 2000US-0226681.								
PR	22-AUG-2000; 2000US-0226868.								
PR	22-AUG-2000; 2000US-0227182.								
PR	23-AUG-2000; 2000US-0227009.								
PR	30-AUG-2000; 2000US-0228924.								
PR	01-SEP-2000; 2000US-0229287.								
PR	01-SEP-2000; 2000US-0229343.								
PR	01-SEP-2000; 2000US-0229344.								
PR	01-SEP-2000; 2000US-0229345.								
PR	05-SEP-2000; 2000US-0229509.								



XX	12-APR-2001 (first entry)	
DT		
XX	S. cerevisiae DNA polymerase protein fragment SEQ ID NO 27.	
DE		
XX	Genome; thermophilic enzyme; washing powder; bleaching.	
KW		
XX	Saccharomyces cerevisiae.	
OS		
XX	WO200075335-A2.	
PN		
XX	14-DEC-2000.	
XX		
XX	02-JUN-2000; 2000WO-IB00893.	
XX		
XX	02-JUN-1999; 99US-0137120.	
XX		
XX	(DECO-) DECODE GENETICS EHF.	
XX		
PI	Hjorleifsdottir S, Hreggvidsson GO, Fridjonsson OH, Aevarsson A;	
PI	Kristjansson JK;	
XX		
XX	WPI; 2001-061727/07.	
DR		
XX	Isolated nucleic acid from bacteriophage RM 378 encoding enzymes useful	
PPT	in recombinant DNA technology -	
PT		
XX	Disclosure; Figure 3A-P; 42pp; English.	
PS		
XX		
CCC	This invention describes a novel isolated nucleic molecule (I) comprising	
CCC	the genome of bacteriophage RM 378. The invention also describes (1) an	
CCC	isolated nucleic acid which encodes a polypeptide obtainable from	
CCC	bacteriophage RM 378, or its active derivative or fragment; (2) an	
CCC	isolated bacteriophage RM 378 (II); (3) a DNA construct (III) comprising	
CCC	operatively linked to a regulatory sequence; (4) a host cell comprising	
CCC	(III); and (5) an isolated polypeptide (IV) obtainable from (II), its	
CCC	active derivative or fragment. Bacteriophage RM 378 is useful for	
CCC	producing thermophilic enzymes useful in DNA research and commercial	
CCC	settings (e.g. proteases and lipases used in washing powder, hydrolytic	
CCC	enzymes used in bleaching). The isolated nucleic acid molecules and	
CCC	vectors are useful in the manufacture of encoded polypeptide, as probes	
CCC	for isolating homologous sequences (e.g. from other bacteriophage	
CCC	species), as well as for detecting the presence of the bacteriophage in	
CCC	a culture of host cells. The polypeptides can be used as a molecular	
CCC	weight marker on SDS-PAGE gels or on molecular sieve gel filtration	
CCC	columns. Because the host organism of the RM378 bacteriophage is a	
CCC	thermophile, the enzymes and proteins of the RM378 bacteriophage are	
CCC	significantly more thermostable than those of other (e.g. mesophilic)	
CCC	bacteriophages, such as the 41 bacteriophage of Escherichia coli. The	
CCC	enhanced stability of the enzymes and proteins of RM378 bacteriophage	
CCC	allows their use under temperature conditions which would be prohibitive	
CCC	for other enzymes, thus increasing the range of conditions which can be	
CCC	employed not only in DNA research but also in commercial settings.	
XX		
SQ	Sequence 761 AA;	
	Query Match 26.7%; Score 55.5; DB 22; Length 761;	
	Best Local Similarity 29.4%; Pred. No.12;	
	Matches 15; Conservative 4; Mismatches 21; Indels 11; Gaps 1	
QY	2 LSNRIILL-----NKAALRKAHTSMVNERFYGRGVQSOLKPRDLG 41	
	:       :    :    :    :	
Db	62 MPNLRCLSLSTQTLMPNPKENQKEIVSTLSAYRNISLDSPENIKPDLG 112	
RESULT 15		
AAV75465		
ID	AAV75465 standard; Protein; 138 AA.	
XX		
AC	AAV75465;	
XX		
DT	21-MAR-2000 (first entry)	
XX		



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 22:02:54 ; Search time 20.0119 Seconds

(without alignments)  
86.686 Million cell updates/sec

Title: US-08-765-244-1

Perfect score: 208

Sequence: 1 MSLNRLILNKALRKAKHTS.....VRNPRYKGPVQSOLKPRDLCL 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents: AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	92	44.2	32	2	US-08-373-190-29 Sequence 29, Appl
2	92	44.2	32	2	US-08-438-190A-29 Sequence 29, Appl
3	92	44.2	32	3	US-08-350-215-29 Sequence 29, Appl
4	92	44.2	32	3	US-09-287-145A-29 Sequence 29, Appl
5	92	44.2	32	4	US-09-556-111-29 Sequence 29, Appl
6	55.5	26.7	761	4	US-09-585-858-27 Sequence 27, Appl
7	52	25.0	457	4	US-09-252-991A-32812 Sequence 32812, A
8	51.5	24.8	208	4	US-09-107-532A-5399 Sequence 5399, Ap
9	50	24.0	310	4	US-09-794-236-4 Sequence 4, Appl
10	50	24.0	593	2	US-08-900-927-1 Sequence 1, Appl
11	50	24.0	593	3	US-09-191-279-1 Sequence 1, Appl
12	50	24.0	593	3	US-09-334-476-1 Sequence 1, Appl
13	49	23.6	547	4	US-09-252-991A-25705 Sequence 25705, A
14	48.5	23.3	274	4	US-09-252-991A-29653 Sequence 29653, A
15	48	23.1	39	1	US-08-428-488-13 Sequence 13, Appl
16	48	23.1	89	4	US-09-732-210-98 Sequence 98, Appl
17	48	23.1	567	3	US-09-188-811-2 Sequence 2, Appl
18	47	22.6	20	5	PCT-US95-07543-5 Sequence 5, Appl
19	47	22.6	91	4	US-09-732-210-195 Sequence 195, App
20	47	22.6	481	4	US-09-130-491-8 Sequence 8, Appl
21	47	22.6	735	3	US-08-539-205A-2 Sequence 2, Appl
22	47	22.6	735	4	US-09-392-163A-2 Sequence 2, Appl
23	47	22.6	905	4	US-09-369-364A-9 Sequence 9, Appl
24	47	22.6	1050	4	US-09-428-711A-16 Sequence 16, Appl
25	47	22.6	2539	3	US-09-413-814-42 Sequence 42, Appl
26	46	22.1	78	2	US-08-327-362-3 Sequence 3, Appl
27	46	22.1	78	3	US-09-158-565-3 Sequence 3, Appl

28	45.5	21.9	399	4	US-09-328-352-7632 Sequence 7632, Ap
29	45.5	21.9	548	3	US-08-903-139B-9 Sequence 9, Appl
30	45.5	21.9	548	3	US-08-903-139B-28 Sequence 28, Appl
31	45	21.6	141	4	US-09-428-711A-4 Sequence 4, Appl
32	45	21.6	218	4	US-09-252-991A-18053 Sequence 18053, A
33	45	21.6	375	4	US-09-428-711A-2 Sequence 2, Appl
34	45	21.6	582	4	US-09-428-711A-2 Sequence 2, Appl
35	45	21.6	609	4	US-09-107-532A-5110 Sequence 5110, Ap
36	45	21.6	730	1	US-07-846-181-5 Sequence 5, Appl
37	45	21.6	730	1	US-07-845-989-5 Sequence 5, Appl
38	45	21.6	795	4	US-09-252-991A-19085 Sequence 19085, A
39	45	21.6	1051	4	US-09-428-711A-14 Sequence 14, Appl
40	45	21.6	1139	4	US-09-914-259-68 Sequence 68, Appl
41	44.5	21.4	119	3	US-09-046-479-4 Sequence 4, Appl
42	44.5	21.4	119	4	US-08-822-897C-4 Sequence 4, Appl
43	44.5	21.4	119	4	US-09-608-810A-5 Sequence 5, Appl
44	44.5	21.4	418	4	US-09-252-991A-25726 Sequence 25726, A
45	44.5	21.4	1164	4	US-09-457-708-2 Sequence 2, Appl

#### ALIGNMENTS

##### RESULT 1

US-08-373-190-29

; Sequence 29, Application US/08373190

; Patent No. 5851829

; GENERAL INFORMATION:

; APPLICANT: MARASCO, WAYNE

; TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF PROTEINS

; NUMBER OF SEQUENCES: 79

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN

; STREET: 130 WATER STREET

; CITY: BOSTON

; STATE: MA

; COUNTRY: US

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/373,190

; FILING DATE: 17-JAN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/06735

; FILING DATE: 16-JUL-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: RESNICK, DAVID S

; REGISTRATION NUMBER: 34,235

; REFERENCE/DOCKET NUMBER: 41956-PCT-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-523-3400

; TELEFAX: 617-523-6440

; TELEX: STRE UR 2002

; INFORMATION FOR SEQ ID NO: 29:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 32 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:

; US-08-373-190-29

Query Match

44.2%; Score 92; DB 2; Length 32;





```

: ORGANISM: Homo sapiens
US-09-794-236-4

Query Match      24.0%; Score 50; DB 4; Length 310;
Best Local Similarity 44.0%; Pred. No. 8.2;
Matches 11; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY      17 AHTSMVRNF--RYGKPVQSOLKPRD 39
        |||||: :| | ||| :|||:
Db      251 AHSILLSFLVRAGKPYDLQIYPQE 275

RESULT 10
US-08-900-927-1
; Sequence 1, Application US/08900927
; Patent No. 5840537
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NEW VESICLE TRANSPORT PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,927
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0350 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: MMLR20T01
; CLONE: 475485
US-08-900-927-1

Query Match      24.0%; Score 50; DB 2; Length 593;
Best Local Similarity 31.8%; Pred. No. 18;
Matches 14; Conservative 9; Mismatches 12; Indels 10; Gaps 2;

QY      4 NLRILNKAALRAKHTSMVRNFRY-----GKPVQSOLKPRD 39
        || | : : ||| :|||:
Db      421 NLAKLIQHANY-QAHSLLRNLRQLGGTVINPGSGTSSRLPRE 464

RESULT 11
US-09-191-279-1
; Sequence 1, Application US/09191279
; Patent No. 5981192
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga

```

```
;
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/191,279
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0350 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: MMLR20T01
; CLONE: 475485
US-09-334-476-1

Query Match      24.0%; Score 50; DB 3; Length 593;
Best Local Similarity 31.1%; Pred. No. 18;
Matches 14; Conservative 9; Mismatches 12; Indels 10

QY      4 NLRILLNKAAALKAKHTSMVRFNY-----GKPVQSOLKPRD 39
           |||::||::||::||::||::||::||::||::||::||::||:
DB      421 NLAKLQHANY-QAHSSLI RNLEQLGGVTNPNGSGTSSRLPRE 464

RESULT 13
US-09-252-991A-25705
; Sequence 25705, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELAT
; TITLE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTIC
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,798
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25705
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25705

Query Match      23.6%; Score 49; DB 4; Length 547;
Best Local Similarity 27.1%; Pred. No. 23;
Matches 13; Conservative 6; Mismatches 7; Indels 22

QY      12 AALRKHAHSMV-----RNFRYKGPVQSOLKP 37
           |||::|||::||::||::||::||::||::||::||:
DB      496 AAAREAHAGMGLDWLRGLKRRTTISNAEIDLRGQAIFYGEPEGR LVP 543

RESULT 14
US-09-252-991A-29653
; Sequence 29653, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELAT
; TITLE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTIC
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

;; PRIOR APPLICATION NUMBER: US 60/074,788  
;; PRIOR FILING DATE: 1998-02-18  
;; PRIOR APPLICATION NUMBER: US 60/094,190  
;; PRIOR FILING DATE: 1998-07-27  
;; NUMBER OF SEQ ID NOS: 33142  
;; SEQ ID NO 29653  
;; LENGTH: 274  
;; TYPE: PRT  
;; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29653

Query Match 23.3%; Score 48.5; DB 4; Length 274;  
Best Local Similarity 46.2%; Pred. NO. 12;  
Matches 12; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

OY 12 AALRKAHTSMVRNFRYKPKVQSLKP 37  
Db 179 AALRTQVDALRNFQOG-PVQAAEQP 203

## RESULT 15

US-08-428-488-13  
; Sequence 13, Application US/08428488  
; Patent No. 5624894  
; GENERAL INFORMATION:  
; APPLICANT: BODOR, Nicholas S.  
; TITLE OF INVENTION: BRAIN-ENHANCED DELIVERY OF NEUROACTIVE  
; TITLE OF INVENTION: PEPTIDES BY SEQUENTIAL METABOLISM  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/428,488  
; FILING DATE: 27-APR-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Mary Katherine  
; REGISTRATION NUMBER: 26,254  
; REFERENCE/DOCKET NUMBER: 028724-087  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 39 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 1  
; OTHER INFORMATION: /note= "Position 1 = H-Ser."  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 39  
; OTHER INFORMATION: /note= "Position 39 = Phe-OH."  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 30  
; OTHER INFORMATION: /note= "Position 30 = Glu-NH2."  
US-08-428-488-13

Query Match 23.1%; Score 48; DB 1; Length 39;  
Best Local Similarity 50.0%; Pred. No. 1.4;  
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 22 VRNFRYKPKVQSLKP 37  
Db 4 MEHFRYKPKVGKKRRP 19

Search completed: July 24, 2003, 22:08:52  
Job time : 21.0119 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 22:02:24 ; Search time 13.6667 Seconds  
(without alignments)  
288.506 Million cell updates/sec

Title: US-08-765-244-1

Perfect score: 208

Sequence: 1 MLNRLILLKALRKRAHTS.....VRNFRYGKPVQSQLKPRDLIC 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR-76:\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	179	86.1	354	1 OWRT	ornithine carbamoy
2	144	69.2	354	1 OWMS	ornithine carbamoy
3	125	60.1	354	1 OHU	ornithine carbamoy
4	112	53.8	41	1 T52779	ornithine transcar
5	57	27.4	350	2 A48421	ornithine transcar
6	57	27.4	351	2 F90403	GTP binding consen
7	56	26.9	1286	2 B71413	hypothetical prote
8	55.5	26.7	1468	2 S58250	DNA-directed DNA p
9	54.5	26.2	138	2 F81900	hypothetical prote
10	53.5	25.7	260	2 G70233	hypothetical prote
11	53.5	25.7	383	2 B83922	short-chain-specif
12	53	25.5	435	2 S31280	cyclin B5 - yeast
13	52.5	25.2	185	2 B81708	translation elonga
14	51.5	24.8	306	2 T52305	En/Spm-like transp
15	51.5	24.8	314	2 F75504	probable pilin, ty
16	51	24.5	240	2 B45359	proopiomelanocorti
17	51	24.5	258	1 CTONEK	corticotropin / li
18	51	24.5	506	2 T50211	WP-repeat protein
19	51	24.5	1779	2 T23130	hypothetical prote
20	49.5	23.8	181	2 T32637	hypothetical prote
21	49.5	23.8	425	2 A40646	sensor-like protei
22	49	23.6	178	2 F91180	hypothetical prote
23	49	23.6	178	2 H86026	hypothetical prote
24	49	23.6	276	2 AD3351	spoU family of rRN
25	49	23.6	300	2 T32702	hypothetical prote
26	49	23.6	320	2 AB0194	pseudouridylylate sy
27	49	23.6	353	2 T25678	hypothetical prote
28	49	23.6	355	1 B69518	GTP-binding protei
29	49	23.6	531	1 NNEC2	anthranilate synth

#### RESULT 1

##### OWRT

ornithine carbamoyltransferase (EC 2.1.1.3) precursor - rat  
N;Alternate names: citrulline phosphorylase; ornithine transcarbamylase  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 28-Feb-1986 #sequence, revision 28-Feb-1986 #text, change 11-Jun-1999  
C;Accession: A00563; A28042; A23090; S02466; I52976; I67609; I53457  
R;Takiguchi, M.; Miura, S.; Mori, M.; Tatibana, M.; Nagata, S.; Kaziro, Y.  
Proc. Natl. Acad. Sci. U.S.A. 81, 7412-7416, 1984  
A;Title: Molecular cloning and nucleotide sequence of cDNA for rat ornithine carbamoyltransferase  
A;Reference number: A00563; MUID:85063800; PMID:6095294  
A;Accession: A00563  
A;Molecule type: mRNA  
A;Residues: 1-354 <TAK1>  
A;Cross-references: GB:K03040; NID:g205873; PIDN:AAA1768.1; PID:g205874  
R;Takiguchi, M.; Murakami, T.; Miura, S.; Mori, M.  
Proc. Natl. Acad. Sci. U.S.A. 84, 6136-6140, 1987  
A;Title: Structure of the rat ornithine carbamoyltransferase gene, a large, X chromosomes  
A;Reference number: A28042; MUID:87317609; PMID:3476935  
A;Accession: A28042  
A;Molecule type: DNA  
A;Residues: 1-354 <TAK2>  
A;Cross-references: GB:M16933; GB:J02957; NID:g205884; PIDN:AAA1769.1; PID:g205886  
R;Kraus, J.P.; Hodges, P.E.; Williamson, C.L.; Horwich, A.L.; Kalousek, F.; Williams, Nucleic Acids Res. 13, 943-952, 1985  
A;Title: A cDNA clone for the precursor of rat mitochondrial ornithine transcarbamylase  
A;Reference number: A23090; MUID:85215524; PMID:3839075  
A;Accession: A23090  
A;Molecule type: mRNA  
A;Residues: 1-38, P', 40-240, S', 242-354 <KRA>  
A;Cross-references: GB:X01976  
R;Aoki, Y.; Sunaga, H.; Suzuki, K.T.  
Biochem. J. 250, 735-742, 1988  
A;Title: A cadmium-binding protein in rat liver identified as ornithine carbamoyltransferase  
A;Reference number: S02466; MUID:88268748; PMID:3390141  
A;Accession: S02466  
A;Molecule type: protein  
A;Residues: 33-56; 293-302; 307-317; 322-329 <AOK>  
R;McIntyre, P.; Graf, L.; Mercer, J.F.B.; Wake, S.A.; Hudson, P.J.; Hoogenraad, N. DNA 4, 147-156, 1985  
A;Title: The primary structure of the imported mitochondrial protein, ornithine transcarbamoyltransferase  
A;Reference number: I52976; MUID:85203360; PMID:3838931  
A;Accession: I52976  
A;Status: translated from GS/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-354 <RES>  
A;Cross-references: GB:M11266; NID:g205871; PIDN:AAA1767.1; PID:g205872  
R;McIntyre, P.; Graf, L.; Mercer, J.; Peterson, G.; Hudson, P.J.; Hoogenraad, N. FEBS Lett. 177, 41-46, 1984  
A;Title: A highly basic N-terminal extension of the mitochondrial matrix enzyme ornithine carbamoyltransferase  
A;Reference number: I53457; MUID:85051832; PMID:6548714  
A;Accession: I67609

#### ALIGNMENTS

30	49	23.6	531	2	C90858	anthranilate synth
31	49	23.6	531	2	E85761	hypothetical prote
32	49	23.6	531	2	AE0653	anthranilate synth
33	49	23.6	538	2	F83354	probable sulfatase
34	49	23.6	986	1	A29646	invasin - Versinia
35	48.5	23.3	333	2	AI0802	Div protein [impor
36	48.5	23.3	945	2	S77052	cation-transportin
37	48.5	23.3	1114	1	S05582	protein-tyrosine k
38	48	23.1	89	2	T43306	ribosomal protein
39	48	23.1	260	2	A71033	probable proteasom
40	48	23.1	260	2	B75181	proteasome endopep
41	48	23.1	362	2	B87613	hypothetical prote
42	48	23.1	633	2	S76749	hypothetical prote
43	48	23.1	765	2	T04186	subtilisin-like pr
44	48	23.1	893	2	T29621	hypothetical prote
45	47.5	22.8	260	2	T11798	cytochrome-c oxida

Matches 33; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 1 MLSNLRLILKAALRKAKHTSMVRFYRGKPVQS--QLKPRDL 40  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1 MLSNRIILNNALKRKGHTSVRFHFCGKPVSQVLKGRDL 42

RESULT 3

OWHU

N:ornithine carbamoyltransferase (EC 2.1.3.3) precursor - human  
C:Species: Homo sapiens (man)  
C>Date: 28-Feb-1996 #sequence\_revision 31-Mar-1993 #text\_change 16-Jun-2000  
C:Accession: A14144; B41444; A00562; I38078; J04672; I59039; I54377  
E:Hata, A.; Tsuzuki, T.; Shimada, K.; Takiguchi, M.; Mori, M.; Matsuda, I.  
J. Biochem. 103, 302-308, 1988  
A:Title: Structure of the human ornithine transcarbamylase gene.  
A:Reference number: A14144; MUID:88227905; PMID:2836378  
A:Accession: A14144  
A:Molecule type: DNA  
A:Residues: 1-354 <HAT>  
A:Cross-references: GB:D00230; NID:g219957; PIDN:BAA00161.1; PID:g219959  
A:Accession: A14144  
A:Molecule type: mRNA  
A:Residues: 1-354 <HA2>  
A:Cross-references: GB:D00230; NID:g219957; PIDN:BAA00161.1; PID:g219959  
R:Horwich, A.L.; Fenton, W.A.; Williams, K.R.; Kalousek, F.; Kraus, J.P.; Doolittle,  
Science 224, 1068-1074, 1984  
A:Title: Structure and expression of a complementary DNA for the nuclear coded precu-  
A:Reference number: A00562; MUID:84196410; PMID:6372096  
A:Accession: A00562  
A:Molecule type: mRNA  
A:Residues: 1-100, 'P', 102-110, 'P', 112-192, 'CF', 195-269, 'R', 271-354 <HOR>  
A:Cross-references: GB:D00230  
R:Hata, A.; Tsuzuki, T.; Shimada, K.; Takiguchi, M.; Mori, M.; Matsuda, I.  
J. Biochem. 100, 717-725, 1986  
A:Title: Isolation and characterization of the human ornithine transcarbamylase ge-  
A:Reference number: I38078; MUID:87057134; PMID:3782067  
A:Accession: I38078  
A>Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-26 <RES>  
A:Cross-references: EMBL:X04443; NID:g35162; PIDN:CAA28039.1; PID:g35163  
R:Wheeler, V.C.; Prodromou, C.P.; Pearl, L.H.; Williamson, R.; Coutelle, C.  
Gene 169, 251-255, 1996  
A:Title: Synthesis of a modified gene encoding human ornithine transcarbamylase for e-  
netic defect.  
A:Reference number: JC4672; MUID:96194812; PMID:8647457  
A:Accession: JC4672  
A:Molecule type: DNA  
A:Residues: 'M', 33-100, 'F', 102-110, 'P', 112-192, 'CF', 195-269, 'R', 271-354 <WHE>  
R:Note: This report represents a synthetic gene designed for expression in (rather th  
R:Horwich, A.L.; Kalousek, F.; Rosenberg, L.E.  
Proc. Natl. Acad. Sci. U.S.A. 82, 4930-4933, 1985  
A:Title: Arginine in the leader peptide is required for both import and proteolytic c  
A:Reference number: I59039; MUID:85270440; PMID:3895227  
A:Accession: I59039  
A>Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-36 <RE2>  
A:Cross-references: GB:M11235; NID:g189408; PIDN:AAA59976.1; PID:g189409  
R:Gilbert-Dussardier, B.; Rabier, D.; Strautnieks, S.; Segues, B.; Bonnefont, J.P.; M  
Hum. Mol. Genet. 3, 831-832, 1994  
A:Title: A novel arginine (245) to glutamine change in exon 8 of the ornithine carba  
C pattern.  
A:Reference number: I54377; MUID:94362689; PMID:8081373  
A:Accession: I54377  
A>Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 269-276, 'Q', 278-289 <RES3>  
A:Cross-references: GB:S73640; NID:g688001; PIDN:AAB31859.1; PID:g688002  
A:Note: This sequence represents a disease defect in ornithine carbamoyltransferase,  
C comment: The active enzyme is a dimeric protein consisting of two identical subunits.



e synthesis of UMP.

C:Comment: The active enzyme catalyzes the condensation of carbamoyl phosphate and ornithine to form citrulline.

C:Genetics:

A:Gene: GDB:OTC

A:Cross-references: GDB:119468; OMIM:311250

A:Map position: Xp21.1-Xp21.1

C:Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase

C:Keywords: arginine biosynthesis; homotrimer; mitochondrion; transferase; urea cycle

F:1-32/Domain: transit peptide (mitochondrion) #status predicted <TNP>

F:33-354/Product: ornithine carbamoyltransferase #status predicted <MAT>

F:40-342/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>

Query Match 60.1%; Score 125; DB 1; Length 354;

Best Local Similarity 66.7%; Pred. No. 7.3e-10;

Matches 28; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY 1 MLSNRLILLNKAALRKAHTSMVRNFRYKGVQSQ--QLKPRDL 40

DB 1 MLENLRLILLNNAAFNGHNFVRNFRGQPLQNKVQLKGRDL 42

RESULT 4

152779

ornithine transcarbamylase peptide - rat (fragment)

C:Species: Rattus sp. (rat)

C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 18-Jun-1999

C:Accession: 152779

R:Horwich, A.L.; Kalousek, F.; Fenton, W.A.; Pollock, R.A.; Rosenberg, L.E.

Cell 44, 451-459, 1986

A:Title: Targeting of pre-ornithine transcarbamylase to mitochondria: Definition of crit

A:Reference number: 152779; MUID:86106223; PMID:3943133

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-41 <RES>

A:Cross-references: GB:M12583; NID:g205887; PIDN:AAA41770.1; PID:g205888

C:Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase

Query Match 53.8%; Score 112; DB 2; Length 41;

Best Local Similarity 61.1%; Pred. No. 4.7e-09;

Matches 22; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MLSNRLILLNKAALRKAHTSMVRNFRYKGVQSQLK 36

DB 1 MLENLRLILLNNAAFNGHNFVRNFRGQPLQNKVQ 36

RESULT 5

A48421

ornithine transcarbamylase - bullfrog

C:Species: Rana catesbeiana (bullfrog)

C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Jun-1999

C:Accession: A48421

R:Helbing, C.; Gergely, G.; Atkinson, B.G.

Dev. Genet. 13, 289-301, 1992

A:Title: Sequential up-regulation of thyroid hormone beta receptor, ornithine transcarba

d thyroid hormone-induced metamorphosis.

A:Reference number: A48421; MUID:93177976; PMID:1291156

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-350 <HEL>

A:Cross-references: GB:M95193; NID:g213683; PIDN:AAA49528.1; PID:g213684

A:Experimental source: liver

A:Note: sequence extracted from NCBI backbone (NCBIN:126154, NCBIP:126155)

C:Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase

C:Keywords: mitochondrion

F:35-338/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>

Query Match

Best Local Similarity 27.4%; Score 57; DB 2; Length 350;

Matches 16; Conservative 9; Mismatches 13; Indels 2; Gaps 2;

QY 1 MLSNRLILLNKAALRKAHTSMVRNFRYKGVQSQLKPRDL 40

DB 1 MLHHMRTIIN-ASWRYGNKICIVRQFGFSQ-TYSOLKGRDL 38

RESULT 6

F90409

GTP binding conserved hypothetical protein SSO3385 [imported] - Sulfolobus solfatar

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001

C:Accession: F90409

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awavez, M.J.; C

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redde

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: F90409

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-351 <KUR>

A:Cross-references: GB:AE006641; NID:g13815687; PIDN:AAK42533.1; GSPDB:GN00155

C:Genetics:

A:Gene: SSO2385

Query Match 27.4%; Score 57; DB 2; Length 351;

Best Local Similarity 32.7%; Pred. No. 3.3;

Matches 17; Conservative 8; Mismatches 9; Indels 18; Gaps 3;

QY 7 ILINKAA-----LRKAHTSMVRNFRY----GKPVQSQ-----LKPRDL 40

DB 294 LILKGGTVLDVARKLHSLAENFRYRVWGKSVKFGQKGVSPSHLEDRDI 345

RESULT 7

B71413

hypothetical protein dl3525w - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: columbia

C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 20-Jun-2000

C:Accession: B71413

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.;

P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weizenecker, T.; Pohl, T.M.; Terry, N.;

avanagh, T.; Hempel, S.; Kotter, P.; Ehtian, K.D.; Rieger, M.; Schaeffer, M.; Funk,

Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdom

erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;

C.; Chalwatzis, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis

A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: B71413

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1286 <BEV>

A:Cross-references: GB:Z97337; NID:g2244829; PID:g2244854

C:Genetics:

A:Map position: 4COP9-4G3845

C:Superfamily: Arabidopsis thaliana hypothetical protein dl3525w

Query Match

Best Local Similarity 26.9%; Score 56; DB 2; Length 1286;

Matches 12; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 3 SNLRILLNKAALRKAHTSMVRNFRYKGVQSQLK 36

DB 107 SNLRNLLSSSTTKRDESLVRNLLVSPILDIO 140

RESULT 8

S58250

DNA-directed DNA polymerase (EC 2.7.7.7) - yeast (Saccharomyces cerevisiae)

N:Alternate names: DNA polymerase I; protein N2181; protein YNL102W

C:Species: Saccharomyces cerevisiae

C:Species: Neisseria meningitidis  
C:date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: F81900  
R:Farkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; McHolroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandram, N.; Rupp, B.; White, O.; Woodward, G.; Barlow, D.N.; Brown, A.J.; Chillingworth, T.P.; Clark, F.; Clifton, H.; Davis, A.P.; de Lencastre, E.; et al.  
Nature 404, 502-506, 2000

A:title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491  
A:Reference number: AB1775; MUID:2022556; PMID:10761919  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-138 <PAR>  
A:Cross-references: GB:AL162755; GB:AL162755; NID:g7379742; PIDN:CAB84568.1; PID:g7379742  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: NMAL316

Query Match 26.2%; Score 54.5; DB 2; Length 138;  
Best Local Similarity 42.9%; Pred. No. 2.7;  
Matches 15; Conservative 4; Mismatches 15; Indels 1; Gaps 1;

QY 6 RILLNKAAALKRKHTSMVNRFRY-GKPVQSOLKPRD 39  
||| ||| ||| ||| ||| |||  
DB 26 RYLLMRRLSTMTATVKLNFRYAGRPKNWLGKYRD 50  
||| ||| ||| ||| ||| |||

RESULT 10  
G70233  
hypothetical protein BBG23 - Lyme disease spirochete plasmid G/ip28-2  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 08-Oct-1999  
C:Accession: G70233  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, R.; Clayton, R.; Lathigra, R.; Wilson, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.; Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:title: Genomic sequence of a Lyme disease spirochaete.  
A:Reference number: A70100; MUID:98065943; PMID:9403685  
A:Accession: G70233  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-260 <KLE>  
A:Cross-references: GB:AE000786; NID:g2690008; PIDN:AAC66071.1; PID:g2690028; TIGR:BEG70233  
A:Experimental source: strain B31  
C:Genetics:  
A:Genome: plasmid

Query Match 25.7%; Score 53.5; DB 2; Length 260;  
Best Local Similarity 21.8%; Pred. No. 7.4;  
Matches 12; Conservative 14; Mismatches 10; Indels 19; Gaps 2;

QY 6 RILLNKAAALKR-----AHTSMVRNFRRYKGPVOS-----QLKPRDLIC 41  
||| :||| ||| :||| ||| :|||  
DB 197 RLVTDYSSIRDSNWSTGSTSFVTHIEYGPIQSWRYPSYTKYSGGDETRPKNTC 251  
||| :||| ||| :||| ||| :|||

RESULT 11  
B83922  
short-chain-specific acyl-CoA dehydrogenase BH2178 [imported] - Bacillus halodurans (C:Species: Bacillus halodurans  
C:date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: B83922  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hara, T.; Nucleic Acids Res. 28, 4317-4331, 2000  
A:title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans A:Reference number: AB3650; MUID:20512582; PMID:11058132  
A:Accession: B83922  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-383 <STO>  
A:Cross-references: GB:AP001514; GB:BA000004; NID:g10174613; PIDN:BAB05897.1; GSPDB:BEG70233  
A:Experimental source: strain C-125



Fri Jul 25 10:14:03 2003

A:Residues: 1-314 <WHI>  
A:Cross-references: GB:AE001913; GB:AE000513; NID:96458240; PIDN:AAF10127.1; PID:9645824  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0548  
A:Map position: 1

Query Match 24.8%; Score 51.5; DB 2; Length 314;  
Best Local Similarity 43.8%; Pred. No. 18;  
Matches 14; Conservative 4; Mismatches 9; Indels 5; Gaps 2;

QY 7 ILINKAALRKRAHTSMVNFNRYGKPVQSOLKPR 38  
Db 85 VLSQKQARPAG---QNERWG-PTQGNVKPR 111

Search completed: July 24, 2003, 22:08:06  
Job time : 16.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 21:59:48 ; Search time 7.32143 seconds  
(without alignments)  
263.350 Million cell updates/sec

Title: US-08-765-244-1

Perfect score: 208

Sequence: 1 MLSNRLILNKAALKAHTS.....VRFNRYKPVQSLKPRDLG 41

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	179	86.1	354	1	OTC_RAT
2	144	69.2	354	1	OTC_MOUSE
3	125	60.1	354	1	OTC_HUMAN
4	57	27.4	350	1	OTC_RANCA
5	55.5	26.7	1468	1	DPOA_YEAST
6	54	26.0	236	1	R82_LOTJA
7	53	25.5	435	1	C8S5_YEAST
8	53	25.5	597	1	IF2P_METAC
9	52.5	25.2	185	1	EFPL_CHLMU
10	51	24.5	226	1	COLI_ONCKE
11	51	24.5	240	1	COLI_ONCKY
12	51	24.5	506	1	POFB_SCHPO
13	50	24.0	591	1	IF2P_METNA
14	50	24.0	593	1	STB2_HUMAN
15	50	24.0	641	1	HRPK_FSESY
16	49	23.6	530	1	TRPG_ECOLI
17	49	23.6	552	1	US1C_HUMAN
18	49	23.6	985	1	INVA_YERPS
19	49	23.6	1739	1	DOTL_HUMAN
20	48.5	23.3	1114	1	REFL_HUMAN
21	48	23.1	88	1	R37A_SCHPO
22	48	23.1	260	1	PSMA_PIRAB
23	48	23.1	260	1	PSMA_PYRHO
24	48	23.1	578	1	SYR_BUCBP
25	47.5	22.8	333	1	Y376_BUCAP
26	47.5	22.8	548	1	NRM1_SHEEP
27	47.5	22.8	906	1	SECA_RICPR
28	47.5	22.8	1110	1	CVGD_BOVIN
29	47	22.6	90	1	R37B_SCHPO
30	47	22.6	260	1	PSMA_PYRFU
31	47	22.6	264	1	UZAG_DROME
32	47	22.6	337	1	MEU1_YEAST
33	47	22.6	748	1	SUF2_HUMAN
34	47	22.6	836	1	YGL1_SCHPO
35	47	22.6	905	1	ATS8_MOUSE
36	47	22.6	1050	1	ULK1_HUMAN
37	46.5	22.4	385	1	PGK_HAEIN
38	46.5	22.4	429	1	TRB2_SULSO
39	46.5	22.4	1475	1	N153_HUMAN
40	46	22.1	177	1	MSRA_SULSO
41	46	22.1	260	1	PSMA_THERK1
42	46	22.1	305	1	FMT_CAMEL
43	46	22.1	333	1	TRPD_YERPE
44	46	22.1	443	1	CGLH_XANNA
45	46	22.1	647	1	PRCA_ANAVA

#### ALIGNMENTS

RESULT 1  
OTC\_RAT  
ID OTC\_RAT STANDARD; PRT; 354 AA.  
AC P00481; Q63407;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)  
DE (OTCase) (Ornithine transcarbamylase).  
GN OTC.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85063800; PubMed=6095294;  
RA Takiguchi M., Miura S., Mori M., Tatibana M., Nagata S., Kaziro Y.;  
RT "Molecular cloning and nucleotide sequence of cDNA for rat ornithine  
carbamoyltransferase precursor.";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:7412-7416(1984).  
RN [2]  
RP SEQUENCE FROM N.A.  
STRAIN=Wistar; TISSUE=Liver;  
RX MEDLINE=87317609; PubMed=3476935;  
RA Takiguchi M., Murakami T., Miura S., Mori M.;  
RT "Structure of the rat ornithine carbamoyltransferase gene, a large, x  
chromosome-linked gene with an atypical promoter.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:6136-6140(1987).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85215524; PubMed=3839075;  
RA Kraus J.P., Hodges P.E., Williamson C.L., Horwich A.L., Kalousek F.,  
PA Williams K.R., Rosenberg L.E.;  
RT "A cDNA clone for the precursor of rat mitochondrial ornithine  
transcarbamylase: comparison of rat and human leader sequences and  
conservation of catalytic sites.";  
RL Nucleic Acids Res. 13:943-952(1985).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85203360; PubMed=3838931;  
RA McIntyre P., Graf L., Mercer J.F.B., Wake S.A., Hudson P.J.,  
RT "The primary structure of the imported mitochondrial protein,  
ornithine transcarbamylase from rat liver: mRNA levels during  
ontogeny.";  
RL DNA 4:147-156(1985).  
RN [5]  
RP SEQUENCE OF 33-56; 293-302; 307-317 AND 321-329.  
RC TISSUE=Liver;  
RX MEDLINE=88266748; PubMed=3390141;  
RA Aoki Y., Sunaga H., Suzuki K.T.;  
RT "A cadmium-binding protein in rat liver identified as ornithine  
carbamoyltransferase.";  
RL Biochem. J. 250:735-742(1988).  
CC -l- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate

Q9y7j8 schizosacch  
P57110 mus musculu  
O75385 homo sapien  
P43726 haemophilus  
Q97tx6 sulfolobus  
P49790 homo sapien  
O97y45 sulfolobus  
O24733 thermococcu  
Q9p728 campylobact  
Q8zeg7 yersinia pe  
P37126 xanthomonas  
P23916 anabaena va

```

CC CC + L-citrulline.
CC CC -|- PATHWAY: SECOND STEP IN UREA CYCLE, ARGININE BIOSYNTHESIS.
CC CC -|- SUBUNIT: Homotrimer.
CC CC -|- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC CC -|- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: K03040; AAA41768.1; -.
CC CC EMBL: M15933; AAA41769.1; -.
CC CC EMBL: M15924; AAA41769.1; JOINED.
CC CC EMBL: M15925; AAA41769.1; JOINED.
CC CC EMBL: M15926; AAA41769.1; JOINED.
CC CC EMBL: M15928; AAA41769.1; JOINED.
CC CC EMBL: M15929; AAA41769.1; JOINED.
CC CC EMBL: M15930; AAA41769.1; JOINED.
CC CC EMBL: M15932; AAA41769.1; JOINED.
CC CC EMBL: X01976; CAA26007.1; -.
CC CC EMBL: K00001; AAA41772.1; -.
CC CC EMBL: M11266; AAA41767.1; -.
CC CC PIR: A00563; OWRT.
CC CC HSP: P00480; IOTH.
CC CC InterPro: IPR006130; Asp/Orn_Cotransf.
CC CC InterPro: IPR002292; Orn_carbtransf.
CC CC InterPro: IPR006131; OTCace_O.
CC CC InterPro: IPR006132; OTCace_P.
CC CC Pfam: PF00185; OTCace_1.
CC CC Pfam: PF02729; OTCace_N; 1.
CC CC PRINTS: PR00100; AOTCASE.
CC CC TIGRfams: TIGR00658; orni_carb.tr; 1.
CC CC PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.
CC CC Arginine biosynthesis; Urea cycle; Transferase; Mitochondrion;
CC CC Transit peptide.
CC CC TRANSIT 1 32 MITOCHONDRION.
CC CC CHAIN 33 354 ORNITHINE CARBAMOYLTRANSFERASE.
CC CC ACT_SITE 263 263 BY SIMILARITY.
CC CC ACT_SITE 303 303 BY SIMILARITY.
CC CC CONFLICT 39 39 G -> P (IN REF. 3; AAA41772).
CC CC CONFLICT 241 241 G -> S (IN REF. 3).
CC CC SEQUENCE 354 AA; 39886 MW; 156B511AF7063F0C CRC64;
CC CC -----
CC CC Query Match 86.1%; Score 179; DB 1; Length 354;
CC CC Best Local Similarity 92.9%; Pred. No. 2.1e-18;
CC CC Matches 39; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
CC CC -----
CC CC QY 1 MLSNRLILNKAALRKAKHTSMVRNFRYKGPVQS--QLKPRDL 40
CC CC ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC CC Db 1 MLSNRLILNKAALRKAKHTSMVRNFRYKGPVQSQVLKGRDL 42
CC CC -----
CC CC RESULT 2
CC CC OTC_MOUSE STANDARD: PRT; 354 AA.
CC CC AC P11725;
CC CC DT 01-OCT-1989 (Rel. 12, Created)
CC CC DT 01-OCT-1989 (Rel. 12, Last sequence update)
CC CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC CC DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)
CC CC (Otcase) (Ornithine transcarbamylase).
CC CC OTC.
CC CC OS Mus musculus (Mouse).
CC CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC CC OX NCBI_TaxID=10090;
CC CC RN [1]
CC CC RP SEQUENCE FROM N.A.
CC CC STRAIN=C57BL/6J;

```

```

RX MEDLINE=87263407; PubMed=3603027;
RA Veres G., Gibbs R.A., Scherer S.E., Caskey C.T.;
RT "The molecular basis of the sparse fur mouse mutation.";
RL Science 237:415-417(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=8815717; PubMed=2831503;
RA Scherer S.E., Veres G., Caskey C.T.;
RT "The genetic structure of mouse ornithine transcarbamylase.";
RL Nucleic Acids Res. 16:1593-1601(1988).
RN [3]
RP SEQUENCE OF 1-26 FROM N.A.
RX STRAIN=C57BL/6J;
RX MEDLINE=86224037; PubMed=3011788;
RA Veres G., Craigen W.J., Caskey C.T.;
RT "The 5' flanking region of the ornithine transcarbamylase gene
contains DNA sequences regulating tissue-specific expression.";
RL J. Biol. Chem. 261:7588-7591(1986).
CC -|- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
+ L-citrulline.
CC -|- PATHWAY: SECOND STEP IN UREA CYCLE, ARGININE BIOSYNTHESIS.
CC -|- SUBUNIT: Homotrimer.
CC -|- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -|- DISEASE: SPARSE FUR (SPF) MOUSE HAVE AN OTC WITH AN OVERALL
DECREASE IN ACTIVITY, AND ALTERED SUBSTRATE AFFINITY.
CC -|- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: M17030; AAA39865.1; -.
CC CC EMBL: M12716; AAA39864.1; ALT_SEQ.
CC CC EMBL: X07092; CAA30121.1; -.
CC CC EMBL: X07093; CAA30121.1; JOINED.
CC CC EMBL: X07094; CAA30121.1; JOINED.
CC CC EMBL: X07095; CAA30121.1; JOINED.
CC CC EMBL: X07096; CAA30121.1; JOINED.
CC CC EMBL: X07097; CAA30121.1; JOINED.
CC CC EMBL: X07098; CAA30121.1; JOINED.
CC CC EMBL: X07099; CAA30121.1; JOINED.
CC CC EMBL: X07100; CAA30121.1; JOINED.
CC CC PIR: A43609; OWMS.
CC CC HSP: P00480; IOTH.
CC CC SWISS-2DPAGE; P11725; MOUSE.
CC CC MGD; MGI:97448; Otc.
CC CC InterPro: IPR006130; Asp/Orn_Cotransf.
CC CC InterPro: IPR002292; Orn_carbtransf.
CC CC InterPro: IPR006131; OTCace_O.
CC CC InterPro: IPR006132; OTCace_P.
CC CC Pfam: PF00185; OTCace_1.
CC CC Pfam: PF02729; OTCace_N; 1.
CC CC PRINTS: PR00100; AOTCASE.
CC CC TIGRfams: TIGR00658; orni_carb.tr; 1.
CC CC PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.
CC CC Arginine biosynthesis; Urea cycle; Transferase; Mitochondrion;
CC CC Transit peptide.
CC CC TRANSIT 1 32 MITOCHONDRION.
CC CC CHAIN 33 354 ORNITHINE CARBAMOYLTRANSFERASE.
CC CC ACT_SITE 263 263 BY SIMILARITY.
CC CC ACT_SITE 303 303 BY SIMILARITY.
CC CC VARIANT 117 117 H -> N (IN SPARSE FUR MOUSE).
CC CC CONFLICT 195 195 G -> R (IN REF. 2).
CC CC SEQUENCE 354 AA; 39765 MW; 33BE5D1E86AA196 CRC64;
CC CC -----
CC CC Query Match 69.2%; Score 144; DB 1; Length 354;
CC CC Best Local Similarity 78.6%; Pred. No. 2.6e-13;
CC CC Matches 33; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

```



RESULT 4  
OTC\_RANCA



**QY** 1 MLSNRLINKAALKRAHTSMVRNFYKGVQSOLKPDL 40  
|| : || : | : | : | : | : | : | : | : | :  
**Dd** 1 MLHMTIIN-ASWRYGNKCIVRQFGSQ-TYSOLKGRDL 38  
| : | : | : | : | : | : | : | : | : | : | :

**RESULT 5**

**DPOA\_YEAST**

**ID DPOA\_YEAST STANDARD; PRT; 1468 AA.**

**AC P13382;**

**DT 01-JAN-1990** (Rel. 13, Created)

**DT 01-OCT-1996** (Rel. 34, Last sequence update)

**DT 28-FEB-2003** (Rel. 41, Last annotation update)

**DE** DNA polymerase alpha catalytic subunit (EC 2.7.7.7) (DNA polymerase I).

**DE GN POL1 OR CDC17 OR YNLJ02W OR N2181.**

**OS Saccharomyces cerevisiae** (Baker's yeast).

**OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.**

**OX NCBI\_TaxID=4932;**

[1]

**RN** SEQUENCE FROM N.A., AND MUTAGENESIS OF GLY-493.  
**RP MEDLINE=68234507; PubMed=3287376;**

**RX Pizzagalli A., Vatsasini P., Plevani P., Luchini G.; "The sequence mapping of a temperature-sensitive mutation, and protein homology with other DNA polymerases." Proc. Natl. Acad. Sci. U.S.A. 85:3772-3776(1988). [2]**

**RN** SEQUENCE FROM N.A.  
**RC STRAIN=S288c// FY1679;**

**RX MEDLINE=96267765; PubMed=8701612;**

**SA** Saiz J.E., Buitrage M.J., Soler A., del Rey F., Revuelta J.L.; "The sequence of a 21.3 kb DNA fragment from the left arm of yeast chromosome XIV reveals LEU4, MET4, POL1, RAS2, and six new open reading frames." Yeast 12:403-409(1996). [3]

**RN** SEQUENCE OF 1-26 FROM N.A.  
**RP STRAIN=S288c;**

**RX MEDLINE=93188702; PubMed=8446029;**

**RA** Mountain H.A., Bystrom A.S., Kotch C.; "The general amino acid control regulates MET4, which encodes a methionine-pathway-specific transcriptional activator of Saccharomyces cerevisiae." Mol. Microbiol. 7:215-228(1993). [4]

**RN** FUNCTION, AND SUBUNIT.  
**RP MEDLINE=20357110; PubMed=10898792;**

**RA** O'H., Zakian V.A.; "The Saccharomyces telomere-binding protein Cdc13p interacts with both the catalytic subunit of DNA polymerase alpha and the telomerase-associated estri protein." Genes Dev. 14:1777-1788(2000).

**CC -!- FUNCTION:** Polymerase alpha in a complex with DNA primase is a replicative polymerase. Has a role in promoting telomere replication during interaction with CDL3.

**CC -!- CATALYTIC ACTIVITY:** N deoxynucleoside triphosphate = N diphosphate + [DNA](N).

**CC -!- SUBUNIT:** Interacts with CDC13.

**CC -!- SUBCELLULAR LOCATION:** Nuclear.

**CC -!- MISCELLANEOUS:** In eukaryotes there are five DNA polymerases: alpha, beta, gamma, delta, and epsilon which are responsible for different reactions of DNA synthesis.

**CC -!- SIMILARITY:** BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch.)

**EMBL; J03268; AAA34888.1; .**  
**EMBL; Z50161; CAAG0524.1; .**  
**EMBL; Z71378; CAA95978.1; .**  
**EMBL; Z12126; CAA78111.1; .**  
**PIR; S58250; S58250.**  
**SBD; S0005046; POL1.**  
**GO; GO:0000731; P:DNA repair synthesis; IMP.**  
**DR InterPro; IPR006172; DNA\_pol\_B.**  
**DR InterPro; IPR006134; DNA\_pol\_B\_dom.**  
**DR InterPro; IPR006133; DNA\_pol\_B\_exo.**  
**DR Pfam; PF00136; DNA\_pol\_B\_1.**  
**DR Pfam; PF03104; DNA\_pol\_B\_exo; 1.**  
**DR PRINTS; PR00106; DNAPOLB.**  
**DR SMART; SM00486; POLBC; 1.**  
**DR TIGRFAMs; TIGR00592; pol2; 1.**  
**DR PROSITE; PS00116; DNA\_POLYMERASE\_B; 1.**  
**KW Transfrase; DNA-directed DNA polymerase; DNA replication;**  
**KW DNA-binding; Nuclear protein.**  
**FT DNA\_BIND 1246..1381 POTENTIAL.**  
**FT VARIANT 493..493 MI -> IV (IN REF. 1).**  
**FT CONFLICT 759..760**  
**SQ SEQUENCE 1468 AA; 166808 MW; 50C9032DBE95BAE CRC64;**

**Query Match 26.7%; Score 55.5; DB 1; Length 1468;**  
**Best Local Similarity 29.4%; Pred. No. 9.1; Indels 11; Gaps 1;**  
**Matches 15; Conservative 4; Mismatches 21;**

**QY 2 LSNNRIILL-----NKALRKHAHTSMVRNFYKGVQSOLKPDL 41**  
: ||| | : : : : : : : : : : : : : : : :  
**Dd 533 MPNLRCLSLSIOTLMNPKNKEIVTSITLSAYRNISLSDSIPENIKPDDL 593**

**RESULT 6**

**RR2\_LOTJA**

**ID RR2\_LOTJA STANDARD; PRT; 236 AA.**

**AC Q9BBS6;**

**DT 28-FEB-2003** (Rel. 41, Created)

**DT 28-FEB-2003** (Rel. 41, Last sequence update)

**DT 28-FEB-2003** (Rel. 41, Last annotation update)

**DE Chloroplast 30S ribosomal protein S2.**

**GN RPS2.**

**OS Lotus japonicus.**

**OG Chloroplast.**

**OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;**

**OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;**

**OC eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.**

**OX NCBI\_taxID=34305;**

[1]

**RN** SEQUENCE FROM N.A.  
**RC STRAIN=Accession PG-20;**  
**RM MEDLIN=21082929; PubMed=11214967;**

**RA Kato T., Kaneko T., Sato S., Nakamura Y., Tabata S.; "Complete structure of the chloroplast genome of a legume, Lotus japonicus";**

**RT** "Lotus Res. 7:323-330(2000)." ;

**CC -!- SIMILARITY:** BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch.)

**EMBL; AP002983; BAB33197.1; .**  
**HMAP; MF\_00291; .. 1.**  
**InterPro; IPR001865; Ribosomal\_S2.**  
**InterPro; IPR005706; S2\_bact\_Org.**  
**Pfam; PF00318; Ribosomal\_S2; 1.**  
**PRINTS; PR00395; RIBOSOMALS2.**  
**TIGRFAMs; TIGR01011; rnsr\_bact; 1**

**WITH CDC28.**

-!- DEVELOPMENTAL STAGE: MAXIMALLY EXPRESSED JUST BEFORE CELL CYCLE START.

-!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN AB SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC PROSITE; PS00962; RIBOSOMAL\_S2\_1; 1.  
DR PROSITE; PS00963; RIBOSOMAL\_S2\_2; 1.  
KW Ribosomal protein; Chloroplast.  
SQ SEQUENCE 236 AA; 26982 MW; CE9E238572322586 CRC64;

Query Match 26.0%; Score 54; DB 1; Length 236;  
Best Local Similarity 38.5%; Pred No. 1.9;  
Matches 10; Conservative 10; Mismatches 4; Indels 2; Gaps 1;

QY 2 LSNRLILLNKA-ALKRAHTSMVRNF 25  
:::||||| : : ||||  
Db 211 IASIRLLNLKLVFAICGHSYSYRNF 236

RESULT 7

CGSS\_YEAST STANDARD; PRT; 435 AA.

AC P30283;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE S-phase entry cyclin 5.  
GN CLB5 OR IPRI20C OR P9642.8.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bf264-15D;  
RX MEDLINE=92387544; PubMed=1387626;  
RA Epstein C.B., Cross F.R.;  
RT \*CLB5: a novel B cyclin from budding yeast with a role in S phase.";  
RN Genes Dev. 6:1695-1706(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB330;  
RX MEDLINE=93307652; PubMed=8319908;  
RA Schwob E., Nasmyth K.;  
RT \*CLB5 and CLB6, a new pair of B cyclins involved in DNA replication  
RN in Saccharomyces cerevisiae";  
RN Genes Dev. 7:1160-1175(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / AB972;  
RX MEDLINE=97313271; PubMed=9169875;  
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W.,  
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,  
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,  
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,  
RA Dietrich F.S., Delius H., Dipalo T., Dubois E., Dueterhoeft A.,  
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,  
RA Hall J., Heblung U., Heumann K., Hilbert H., Hillier L.,  
RA Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,  
RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,  
RA Marathe R., Messinguy F., Newes H.-W., Miripati S., Moestl D.,  
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,  
RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Schafie M.,  
RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tetelin H.,  
RA Urrestarazu L.A., Ushinsky S., Vierendeels F., Viessers S., Voss H.,  
RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,  
RA Zhong W.W., Zollner A., Vo D.H., Hani J.;  
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI";  
RN Nature 387:103-105(1997).  
RL Nature 387:103-105(1997).

-!- FUNCTION: REQUIRED FOR EFFICIENT PROGRESSION THROUGH S PHASE AND  
CC POSSIBLY FOR THE NORMAL PROGRESSION THROUGH MEIOSIS. INTERACTS

CC WITH CDC28.

-!- DEVELOPMENTAL STAGE: MAXIMALLY EXPRESSED JUST BEFORE CELL CYCLE START.

-!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN AB SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC PROSITE; PS00962; RIBOSOMAL\_S2\_1; 1.  
DR PROSITE; PS00963; RIBOSOMAL\_S2\_2; 1.  
KW Ribosomal protein; Chloroplast.  
SQ SEQUENCE 236 AA; 26982 MW; CE9E238572322586 CRC64;

Query Match 26.0%; Score 54; DB 1; Length 236;  
Best Local Similarity 38.5%; Pred No. 1.9;  
Matches 10; Conservative 10; Mismatches 4; Indels 2; Gaps 1;

QY 2 LSNRLILLNKA-ALKRAHTSMVRNF 25  
:::||||| : : ||||  
Db 211 IASIRLLNLKLVFAICGHSYSYRNF 236

RESULT 7

CGSS\_YEAST STANDARD; PRT; 435 AA.

AC P30283;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE S-phase entry cyclin 5.  
GN CLB5 OR IPRI20C OR P9642.8.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bf264-15D;  
RX MEDLINE=92387544; PubMed=1387626;  
RA Epstein C.B., Cross F.R.;  
RT \*CLB5: a novel B cyclin from budding yeast with a role in S phase.";  
RN Genes Dev. 6:1695-1706(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB330;  
RX MEDLINE=93307652; PubMed=8319908;  
RA Schwob E., Nasmyth K.;  
RT \*CLB5 and CLB6, a new pair of B cyclins involved in DNA replication  
RN in Saccharomyces cerevisiae";  
RN Genes Dev. 7:1160-1175(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / AB972;  
RX MEDLINE=97313271; PubMed=9169875;  
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W.,  
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,  
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,  
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,  
RA Dietrich F.S., Delius H., Dipalo T., Dubois E., Dueterhoeft A.,  
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,  
RA Hall J., Heblung U., Heumann K., Hilbert H., Hillier L.,  
RA Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,  
RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,  
RA Marathe R., Messinguy F., Newes H.-W., Miripati S., Moestl D.,  
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,  
RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Schafie M.,  
RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tetelin H.,  
RA Urrestarazu L.A., Ushinsky S., Vierendeels F., Viessers S., Voss H.,  
RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,  
RA Zhong W.W., Zollner A., Vo D.H., Hani J.;  
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI";  
RN Nature 387:103-105(1997).  
RL Nature 387:103-105(1997).

-!- FUNCTION: REQUIRED FOR EFFICIENT PROGRESSION THROUGH S PHASE AND  
CC POSSIBLY FOR THE NORMAL PROGRESSION THROUGH MEIOSIS. INTERACTS

CC WITH CDC28.

-!- DEVELOPMENTAL STAGE: MAXIMALLY EXPRESSED JUST BEFORE CELL CYCLE START.

-!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN AB SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC PROSITE; PS00962; RIBOSOMAL\_S2\_1; 1.  
DR PROSITE; PS00963; RIBOSOMAL\_S2\_2; 1.  
KW Ribosomal protein; Chloroplast.  
SQ SEQUENCE 236 AA; 26982 MW; CE9E238572322586 CRC64;

Query Match 26.0%; Score 54; DB 1; Length 236;  
Best Local Similarity 38.5%; Pred No. 1.9;  
Matches 10; Conservative 10; Mismatches 4; Indels 2; Gaps 1;

QY 2 LSNRLILLNKA-ALKRAHTSMVRNF 25  
:::||||| : : ||||  
Db 211 IASIRLLNLKLVFAICGHSYSYRNF 236

RESULT 7

CGSS\_YEAST STANDARD; PRT; 435 AA.

AC P30283;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE S-phase entry cyclin 5.  
GN CLB5 OR IPRI20C OR P9642.8.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bf264-15D;  
RX MEDLINE=92387544; PubMed=1387626;  
RA Epstein C.B., Cross F.R.;  
RT \*CLB5: a novel B cyclin from budding yeast with a role in S phase.";  
RN Genes Dev. 6:1695-1706(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB330;  
RX MEDLINE=93307652; PubMed=8319908;  
RA Schwob E., Nasmyth K.;  
RT \*CLB5 and CLB6, a new pair of B cyclins involved in DNA replication  
RN in Saccharomyces cerevisiae";  
RN Genes Dev. 7:1160-1175(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / AB972;  
RX MEDLINE=97313271; PubMed=9169875;  
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W.,  
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,  
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,  
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,  
RA Dietrich F.S., Delius H., Dipalo T., Dubois E., Dueterhoeft A.,  
RA Duncan M., Floeth M., Fort

CC the binding of the formylmethionine-tRNA to ribosomes. Seems to  
 CC function along with eIF-2 (by similarity).  
 CC -1- SIMILARITY: BELONGS TO THE IF-2 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AE010822; AAM04939.1; -  
 CC HAMAP; MF\_00100; -; 1.  
 CC InterPro; IPR000795; EF\_GTPbind.  
 CC InterPro; IPR004163; EFTU\_D2.  
 CC InterPro; IPR000178; If2.  
 CC InterPro; IPR005225; Small\_GTP.  
 CC InterPro; IPR004544; TIF\_aif-2.  
 CC Pfam; PF00009; GTP\_EFTU; 1.  
 CC Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 CC TIGRFAMS; TIGR00491; aIF-2; 1.  
 CC TIGRFAMS; TIGR00231; Small\_GTP; 1.  
 CC PROSITE; PS01176; If2; FALSE\_NEG.  
 CC Initiation factor; Protein biosynthesis; GTP-binding;  
 CC Complete proteome.  
 CC NP\_BIND 22 29 GTP (BY SIMILARITY).  
 CC NP\_BIND 84 88 GTP (BY SIMILARITY).  
 CC NP\_BIND 138 141 GTP (BY SIMILARITY).  
 CC SEQUENCE 597 AA; 65438 MW; 67CAF4D902C1B8D5 CRC64;  
 CC -----  
 CC Query Match 25.5%; Score 53; DB 1; Length 597;  
 CC Best Local Similarity 31.0%; Pred. No. 7.7;  
 CC Matches 13; Conservative 12; Mismatches 11; Indels 6; Gaps 2;  
 CC -----  
 CC QY 3 SNLRLLKALRKAHTSMVNRFGKPVQSQ----LKPRDL 40  
 CC Db 255 ATLDVILDTLKKGDTWVIGSL-GEPIQTKVALLKPREL 294  
 CC -----  
 CC RESULT 9  
 CC EFP1\_CHLMU STANDARD; PRT; 185 AA.  
 CC AC QPQR6;  
 CC DT 16-OCT-2001 (Rel. 40, Created)  
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Elongation factor P 1 (EF-p 1).  
 CC GN EFP1 OR TC0398.  
 CC OS Chlamydia muridarum.  
 CC OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
 CC OX NCBI\_TaxID=83560;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN-MoPn / Ni9g;  
 CC RX MEDLINE=20150255; PubMed=10684935;  
 CC RA White T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
 CC RA Read O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,  
 CC RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
 CC RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,  
 CC RA Eisen J., Fraser C.M.;  
 CC RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia  
 CC RT pneumoniae AR39";  
 CC RL Nucleic Acids Res. 28:1397-1406(2000).  
 CC CC -1- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATES EFFICIENT  
 CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED  
 CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING  
 CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-tRNA, THUS INCREASING  
 CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE  
 CC (BY SIMILARITY).  
 CC CC -1- PATHWAY: Protein biosynthesis.  
 CC CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC CC -1- SIMILARITY: Belongs to the elongation factor P family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AE002306; AAF39255.1; -  
 CC FIR; B81708; B81708.  
 CC TIGR; TC0398; -  
 CC HAMAP; MF\_00141; -; 1.  
 CC InterPro; IPR001059; EF-P.  
 CC Pfam; PF01132; EFP; 1.  
 CC PROSITE; PS01275; EFP; 1.  
 CC Protein biosynthesis; Elongation factor; Complete proteome.  
 CC SEQUENCE 185 AA; 20468 MW; 969BD834F4997CE2 CRC64;  
 CC -----  
 CC Query Match 25.2%; Score 52.5; DB 1; Length 185;  
 CC Best Local Similarity 42.4%; Pred. No. 2.4;  
 CC Matches 14; Conservative 9; Mismatches 7; Indels 3; Gaps 2;  
 CC -----  
 CC QY 11 KAALRKAHTSMV--RNFYRGKPV-QSQLKPRDL 40  
 CC Db 38 KVALQAAGSDVVVVERNFKAGQEVKEAQPRL 70  
 CC -----  
 CC RESULT 10  
 CC COLI\_ONCRE STANDARD; PRT; 226 AA.  
 CC AC P10000; P01199; P01204; P87470; P87471; P87472; P87473; P87474;  
 CC AC P87475; P87476; P87477; P87478; P87479; Q90521; Q92024;  
 CC DT 21-JUL-1986 (Rel. 01, Created)  
 CC DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Corticotropin-lipotropin precursor (Pro-opiomelanocortin) (POMC)  
 CC [Contains: Corticotropin (Adrenocorticotrophic hormone) (ACTH);  
 CC Melanotropin alpha (Alpha-MSH); Corticotropin-like intermediate  
 CC peptide (CLIP); Lipotropin beta (Beta-LPH); Lipotropin gamma (Gamma-  
 CC LPH); Melanotropin beta (Beta-MSH); Beta-endorphin; Met-enkephalin].  
 CC OS Oncorhynchus keta (Chum salmon).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 CC OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 CC OX NCBI\_TaxID=8018;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=89064064; PubMed=3197404;  
 CC RA Kitahara N., Nishizawa T., Iida K., Okazaki H., Andoh T., Soma G.-I.;  
 CC RT "Absence of a gamma-melanocyte-stimulating hormone sequence in  
 CC proopiomelanocortin mRNA of chum salmon *Oncorhynchus keta*.";  
 CC RL Comp. Biochem. Physiol. 91B:365-370(1988).  
 CC RN [2]  
 CC RP SEQUENCE OF 92-226 FROM N.A.  
 CC RX MEDLINE=85062796; PubMed=6095185;  
 CC RA Soma G.-I., Kitahara N., Nishizawa T., Nanami H., Kotake C.,  
 CC RA Okazaki H., Andoh T.;  
 CC RT "Nucleotide sequence of a cloned cDNA for proopiomelanocortin  
 CC precursor of chum salmon, *Oncorhynchus keta*.";  
 CC RL Nucleic Acids Res. 12:8029-8041(1984).  
 CC RN [3]  
 CC RP SEQUENCE OF 92-226 FROM N.A.  
 CC RX MEDLINE=84280046; PubMed=6087806;  
 CC RA Nishizawa T., Kitahara N., Nanami H., Hara N., Kotake C., Okazaki H.,  
 CC RA Andoh T., Soma G.-I.;  
 CC RT "Heterogeneity of 3' nontranslated regions in proopiomelanocortin  
 CC (POMC) precursor mRNA of chum salmon *Oncorhynchus keta*: polymorphism  
 CC of the gene".  
 CC RL Biochem. Biophys. Res. Commun. 122:556-562(1984).  
 CC RN [4]  
 CC RP SEQUENCE OF 98-112.  
 CC RX MEDLINE=81086168; PubMed=7447938;

RA Kawauchi H., Adachi Y., Tsubokawa M.;  
 RT "Occurrence of a new melanocyte stimulating hormone in the salmon  
 RL pituitary gland.";  
 RL Biochem. Biophys. Res. Commun. 96:1508-1517(1980).  
 RN [5]  
 RP SEQUENCE OF 198-226.  
 RX MEDLINE=79255548; PubMed=475783;  
 RA Kawauchi H., Tsubokawa M., Muramoto K.;  
 RT "Isolation and primary structure of endorphin from salmon pituitary  
 RL glands.";  
 RL Biochem. Biophys. Res. Commun. 88:1249-1254(1979).  
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES AT PAIRED BASIC RESIDUES YIELD  
 CC THE DIFFERENT ACTIVE PEPTIDES.  
 CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.  
 CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; M27692; AAA49426.1; ALT\_INIT.  
 DR EMBL; K02613; AAA49424.1; .  
 DR EMBL; K02614; AAA49425.1; .  
 DR EMBL; X01122; CAA25591.1; .  
 DR PIR; I51080; CTONPK.  
 DR InterPro; IPR001941; Mcoirtin\_ACTH.  
 DR Pfam; PF00976; ACTH\_domain; 1.  
 DR PRINTS; PR00383; MELANOCORTIN.  
 DR ProDom; PD003250; Mcoirtin\_ACTH; 1.  
 DR Endorphin; Hormone; Cleavage on pair of basic residues; Acetylation;  
 KW Signal.  
 FT SIGNAL 1 22 POTENTIAL.  
 FT PEPTIDE 98 138 CORTICOTROPIN.  
 FT PEPTIDE 98 112 MELANOTROPIN ALPHA.  
 FT PEPTIDE 116 138 CORTICOTROPIN-LIKE INTERMEDIARY PEPTIDE.  
 FT PEPTIDE 141 226 LIPOTROPIN BETA.  
 FT PEPTIDE 141 195 LIPOTROPIN GAMMA.  
 FT PEPTIDE 179 195 LIPOTROPIN GAMMA.  
 FT PEPTIDE 179 195 MELANOTROPIN BETA.  
 FT PEPTIDE 198 226 BETA-ENDORPHIN.  
 FT PEPTIDE 198 98 MET-ENKEPHALIN.  
 FT MOD\_RES 98 98 ACETYLATION.  
 SQ SEQUENCE 226 AA; 24982 MW; 327CA785F69B1B24 CRC64;  
 Query Match 24.5%; Score 51; DB 1; Length 226;  
 Best Local Similarity 36.4%; Pred. No. 5;  
 Matches 8; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
 QY 16 KAHTSMVRNFRYKPKVQSQLKP 37  
 ID COLJ\_ONCMY STANDARD; PRT; 240 AA.  
 DB 95 KRHSYSMEHFRWGKPIGHKRRP 116  
 ID COLJ\_ONCMY STANDARD; PRT; 240 AA.  
 AC Q04618;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Corticotropin-lipotropin B precursor (Pro-opiomelanocortin) (POMC).  
 GN POMC-B.  
 OS Onchorynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE= Pituitary;  
 RX MEDLINE=93078794; PubMed=1448114;

RA Salbert G., Chauveau I., Bonnet G., Valotaire Y., Jégo P.;  
 RT "One of the two trout proopiomelanocortin messenger RNAs potentially  
 RL encodes new peptides.";  
 RL Mol. Endocrinol. 6:1605-1613(1992).  
 CC -1- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS OF ADULT DIPLOID  
 CC ANIMALS.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED ONLY IN SEXUALLY ACTIVE FISH.  
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES AT PAIRED BASIC RESIDUES YIELD  
 CC THE DIFFERENT ACTIVE PEPTIDES.  
 CC -1- PTM: ACETYLATION OF BETA-ENDORPHIN OCCURS IN A TISSUE-SPECIFIC  
 CC MANNER.  
 CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.  
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-15 IS THE INITIATOR.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; X69809; CAA49467.1; .  
 DR PIR; B45359; B45359.  
 DR InterPro; IPR001941; Mcoirtin\_ACTH.  
 DR Pfam; PF00976; ACTH\_domain; 1.  
 DR ProDom; PD003250; Mcoirtin\_ACTH; 1.  
 DR Endorphin; Hormone; Cleavage on pair of basic residues; Amidation;  
 KW Signal; Acetylation; Pyrrolidone carboxylic acid.  
 FT SIGNAL 1 36 POTENTIAL.  
 FT PEPTIDE 37 108 NPP 2 (BY SIMILARITY).  
 FT PEPTIDE 112 152 CORTICOTROPIN (ACTH) (BY SIMILARITY).  
 FT PEPTIDE 112 124 MELANOTROPIN ALPHA 2 (ALPHA-MSH 2)  
 FT PEPTIDE 130 152 (BY SIMILARITY).  
 FT PEPTIDE 130 152 CORTICOTROPIN-LIKE INTERMEDIARY PEPTIDE 2  
 FT PEPTIDE 155 240 (CLIP 2) (BY SIMILARITY).  
 FT PEPTIDE 155 240 LIPOTROPIN BETA (LPH-BETA)  
 FT PEPTIDE 155 209 (BY SIMILARITY).  
 FT PEPTIDE 193 209 LIPOTROPIN GAMMA (LPH-GAMMA)  
 FT PEPTIDE 193 209 (BY SIMILARITY).  
 FT PEPTIDE 212 240 MELANOTROPIN BETA 2 (BETA-MSH 2)  
 FT PEPTIDE 212 216 BETA-ENDORPHIN 1 (BY SIMILARITY).  
 FT MOD\_RES 37 37 MET-ENKEPHALIN.  
 FT MOD\_RES 112 112 PYRROLIDONE CARBOXYLIC ACID  
 FT MOD\_RES 124 124 (BY SIMILARITY).  
 FT MOD\_RES 124 124 ACETYLATION (BY SIMILARITY).  
 FT MOD\_RES 124 124 AMIDATION (G-125 PROVIDE AMIDE GROUP)  
 FT MOD\_RES 124 124 (BY SIMILARITY).  
 SQ SEQUENCE 240 AA; 26719 MW; 4F715CE8E6424F6C CRC64;  
 Query Match 24.5%; Score 51; DB 1; Length 240;  
 Best Local Similarity 36.4%; Pred. No. 5.4;  
 Matches 8; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
 QY 16 KAHTSMVRNFRYKPKVQSQLKP 37  
 ID COLJ\_ONCMY STANDARD; PRT; 240 AA.  
 DB 109 KRHSYSMEHFRWGKPIGHKRRP 130  
 ID COLJ\_ONCMY STANDARD; PRT; 240 AA.  
 AC Q09855; Q9P7V1;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein pofll.  
 GN POF11 OR SPAC29E6.01 OR SPAC30.05.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.



```
Best Local Similarity 31.0%; Pred. No. 21;
Matches 13; Conservative 12; Mismatches 11; Indels 6; Gaps 2;

QY 3 SNLRLLINKAALRAKHTSMVNFYRGKPVQSQ-----LKPRDL 40
DB 249 ATLDLIIYDGLTKAGDTWIGSL--GPIRTKVRALLKREL 288

RESULT 14
STB2_HUMAN
ID STB2_HUMAN STANDARD; PRT; 593 AA.
AC Q15833;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Syntaxin binding protein 2 (Unc-18 homolog 2) (Unc-18B) (Unc18-2).
GN STXB2P2 OR UNC18B.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97079654; PubMed=8921365;
RA Ziegler S.F., Mortrud M.T., Swartz A.R., Baker E., Sutherland G.R.,
RA Burmeister M., Mulligan J.T.;
RT "Molecular characterization of a nonneuronal human UNC18 homolog.";
RL Genomics 37:19-23(1996).
CC -1- FUNCTION: INVOLVED IN THE PROTEIN TRAFFICKING FROM THE GOLGI
CC APPARATUS TO THE PLASMA MEMBRANE. BINDS SYNTAXINS 1A, 2, 3 BUT NOT
CC 4 (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: PLACENTA, LUNG, LIVER, KIDNEY AND PANCREAS, AS
CC WELL AS IN PERIPHERAL BLOOD LYMPHOCYTES.
CC -1- SIMILARITY: BELONGS TO THE STXB2P/UNC-18/SECL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: U63533; AAC50762.1;
Genew; HGNC:11445; STXB2P.
MIM; 601717;
DR GO: GO:0030347; F:syntaxin-2 binding activity; NAS.
DR GO: GO:0030348; F:syntaxin-3 binding activity; NAS.
DR GO: GO:0006886; P:intracellular protein transport; NAS.
DR InterPro: IPR001619; Secl-like.
DR Pfam: PF00995; Secl; 1.
KW Protein transport; Polymorphism.
FT VARIANT 526 526 /FTID=VAR_014934.
SQ SEQUENCE 593 AA; 66438 MW; 5BE27B47B2017A2B CRC64;

Query Match 24.0%; Score 50; DB 1; Length 593;
Best Local Similarity 31.1%; Pred. No. 21;
Matches 14; Conservative 9; Mismatches 12; Indels 10; Gaps 2;

QY 4 NLRLLINKAALRAKHTSMVNFYRGKPVQSQ-----GKPVQSQKPRD 39
DB 421 NLAKLIQHNV-QAHSLLRNLEQLGGVTNPGSGTSSRLPRE 464

RESULT 15
HRPK_PSEY
ID HRPK_PSEY STANDARD; PRT; 641 AA.
AC P41501;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Pathogenicity locus protein hrpk.
OS Pseudomonas syringae (pv. syringae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=321;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PSS61;
RX MEDLINE=94148760; PubMed=8106313;
RA Xiao Y., Heu S., Yi J., Lu Y., Hutcheson S.W.;
RT "Identification of a putative alternate sigma factor and
RT characterization of a multicomponent regulatory cascade controlling
RT the expression of Pseudomonas syringae pv. syringae Pss61 hrp and
RT hrpA genes.";
RL J. Bacteriol. 176:1025-1036(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=PSS61;
RX MEDLINE=941100578; PubMed=8274770;
RA Heu S., Hutcheson S.W.;
RT "Nucleotide sequence and properties of the hrpA locus associated with
RT the Pseudomonas syringae pv. syringae 61 hrp gene cluster.";
RL Mol. Plant Microbe Interact. 6:553-564(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: U03855; AAA17653.1;
SEQUENCE 641 AA; 67678 MW; EC098941E5B46C8E CRC64;

Query Match 24.0%; Score 50; DB 1; Length 641;
Best Local Similarity 33.3%; Pred. No. 23;
Matches 11; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 7 ILLNKAAALRAKHTSMVNFYRGKPVQSQKPRD 39
DB 447 LVRSLSALQKAVTEQAQNVNSGKALQTDLATD 479

Search completed: July 24, 2003, 22:04:31
Job time : 10.3214 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 22:01:38 ; Search time 29.2857 seconds  
(without alignments)  
361.274 Million cell updates/sec

Title: US-08-765-244-1

Perfect score: 208

Sequence: 1 MLSNRLINKAALRAKHAHTS.....VNRFRYGRPVQSOLKPRDL 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	179	86.1	102	11	P70518	P70518 rattus norv
2	144	69.2	351	11	Q8RIA8	Q8RIA8 mus musculu
3	114	54.8	354	6	Q9N1U7	Q9N1U7 bos taurus
4	112	53.8	41	11	Q63786	Q63786 rattus norv
5	92	44.2	354	13	Q9IAU8	Q9IAU8 trachemys s
6	57	27.4	351	17	Q97W55	Q97W55 sulfolobus
7	56	26.9	356	13	Q9IAV0	Q9IAV0 sceloporos
8	56	26.9	1286	10	Q23351	Q23351 arabidopsis
9	54.5	26.2	138	16	Q9JUG9	Q9JUG9 neisseria m
10	54	26.0	457	13	Q8AVS9	Q8AVS9 xenopus lae
11	54	26.0	544	5	Q56570	Q56570 drosophila
12	54	26.0	769	2	Q52393	Q52393 pseudomonas
13	53.5	25.7	260	16	Q50746	Q50746 borrelia bu
14	53.5	25.7	383	16	Q9KAV8	Q9KAV8 bacillus ha
15	53	25.5	499	2	Q8KNV7	Q8KNV7 clostridium
16	52.5	25.2	140	12	Q99A14	Q99A14 bovine vira

17	52.5	25.2	141	16	Q8DG64	Q8dg64 synechococ
18	52.5	25.2	475	5	Q9NGB5	Q9ngb5 drosophila
19	52.5	25.2	942	5	Q95VA4	Q95va4 drosophila
20	52.5	25.2	1138	5	Q95VA3	Q95va3 drosophila
21	52.5	25.2	1608	5	Q95VA5	Q95va5 drosophila
22	51.5	24.8	306	10	Q9S728	Q9s728 arabidopsis
23	51.5	24.8	314	16	Q9RWV7	Q9rww7 deinococcus
24	51.5	24.8	346	2	Q93UJ3	Q93uj3 burkholderi
25	51.5	24.8	346	2	Q9AI37	Q9ai37 burkholderi
26	51	24.5	1779	5	O18150	O18150 caenorhabdi
27	50.5	24.3	296	5	Q9VXZ7	Q9vxz7 drosophila
28	50.5	24.3	755	5	Q9V4J5	Q9v4j5 drosophila
29	50.5	24.3	781	5	Q8IGP7	Q8igp7 drosophila
30	50	24.0	191	12	Q8JLA0	Q8jla0 ectromelia
31	50	24.0	211	2	Q9R6K6	Q9r6k6 agrobacteri
32	50	24.0	310	4	Q9EBM4	Q9ebm4 homo sapien
33	50	24.0	318	4	Q96B22	Q96b22 homo sapien
34	50	24.0	376	5	Q8SUE7	Q8sue7 encephalito
35	50	24.0	465	4	Q9HBM3	Q9hbm3 homo sapien
36	50	24.0	579	4	O00184	O00184 homo sapien
37	50	24.0	590	4	Q9BU65	Q9bu65 homo sapien
38	50	24.0	632	4	Q96JX1	Q96jx1 homo sapien
39	50	24.0	654	4	Q9H6X0	Q9h6x0 homo sapien
40	50	24.0	702	4	Q8IZ21	Q8iz21 homo sapien
41	50	24.0	787	12	Q83944	Q83944 olive laten
42	50	24.0	831	4	Q8NEM5	Q8nem5 homo sapien
43	50	24.0	866	5	Q9GPC3	Q9gpg3 drosophila
44	50	24.0	866	5	Q9W034	Q9w0e4 drosophila
45	50	24.0	882	4	Q9HBM5	Q9hbm5 homo sapien

## ALIGNMENTS

### RESULT 1

P70518 PRELIMINARY; PRT; 102 AA.  
AC P70518;  
DT 01-FEB-1997 (TRENBLrel. 02, Created)  
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Ornithine carbamoyltransferase precursor (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85051832; PubMed=6548714;  
RA McIntyre P., Graf L., Mercer J., Peterson G., Hudson P.J.,  
RA Hoogenraad N.;  
RT "A highly basic N-terminal extension of the mitochondrial matrix  
enzyme ornithine transcarbamylase from rat liver.";  
RL FEBS Lett. 177:41-46(1984).  
DR EMBL; K03041; AAA41771.1; -.  
DR HSP; P00480; 10TH.  
DR InterPro; IPR006130; Asp/Orn\_Cotranf.  
DR InterPro; IPR006132; OTCace\_P.  
DR Pfam; PF02729; OTCace\_N; 1.  
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
FT SIGNAL 1 32 POTENTIAL.  
FT CHAIN 33 >102 ORNITHINE CARBAMOYLTRANSFERASE.  
FT NON\_TER 102 102  
SQ SEQUENCE 102 AA; 11932 MW; B739C41F19BF720D CRC64;

Query Match 86.1%; Score 179; DB 11; Length 102;  
Best Local Similarity 92.9%; Pred. No. 5.3e-19;  
Matches 39; Conservative 0; Mismatches 1; Indels 2; Gaps 1;  
QY 1 MLSNRLINKAALRAKHAHTSWNRFRYGRPVQS--OLKPRDL 40  
|||||  
Db 1 MLSNRLINKAALRAKHAHTSWNRFRYGRPVQSOLKGRDL 42

## RESULT 2

Q8RLA8 ID Q8RLA8 PRELIMINARY; PRT; 351 AA.  
 AC Q8RLA8;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Similar to ornithine transcarbamylase.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Scrausberg A.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; BC024893; AAH24893.1; -  
 DR HSSP; P00479; 3CSU.  
 DR InterPro; IPR006130; Asp/Orn\_Cotranf.  
 DR InterPro; IPR002292; Orn\_carbtransf.  
 DR InterPro; IPR006131; OTCace\_O.  
 DR InterPro; IPR006132; OTCace\_P.  
 DR Pfam; PF00185; OTCace; 1.  
 DR Pfam; PF02729; OTCace.N; 1.  
 DR TIGRFAMs; TIGR00658; orni\_carb\_tr; 1.  
 DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
 SQ SEQUENCE 351 AA; 39365 MW; E6B38B2FC779F42 CRC64;

Query Match 69.2%; Score 144; DB 11; Length 351;

Best Local Similarity 78.6%; Pred. No. 3.9e-13;

Matches 33; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 1 MLSNRLILLNKAALRKAHTSMVRNFRYKGPVQS--QLKPRDL 40  
 |||||  
 DB 1 MLSNRLILLNKAALRKAHTSMVRNFRYKGPVQS--QLKPRDL 40

## RESULT 3

Q9NLU7 ID Q9NLU7 PRELIMINARY; PRT; 354 AA.  
 AC Q9NLU7;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Ornithine transcarbamylase.  
 GN OTC.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shimogiri T., Koyanagi K., Mannen H., Tsuji S.;  
 RT "Ornithine transcarbamylase is expressed in uricotelic animals.";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF134841; AAF61405.1; -  
 DR HSSP; P00480; 1OTH.  
 DR InterPro; IPR006130; Asp/Orn\_Cotranf.  
 DR InterPro; IPR002292; Orn\_carbtransf.  
 DR InterPro; IPR006131; OTCace\_O.  
 DR InterPro; IPR006132; OTCace\_P.  
 DR Pfam; PF00185; OTCace; 1.  
 DR Pfam; PF02729; OTCace.N; 1.  
 DR PRINTS; PR00100; AOTCASE.  
 DR TIGRFAMs; TIGR00658; orni\_carb\_tr; 1.  
 DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
 SQ SEQUENCE 354 AA; 39842 MW; 30EAC2EF4AC1D71B CRC64;

Query Match 54.8%; Score 114; DB 6; Length 354;

Best Local Similarity 61.9%; Pred. No. 1.3e-08;  
 Matches 26; Conservative 4; Mismatches 10; Indels 2; Gaps 1;

QY 1 MLSNRLILLNKAALRKAHTSMVRNFRYKGPVQS--QLKPRDL 40  
 |||||  
 DB 1 MLFHLRTLLNNAALRNHGFVVRNFRGCGPLQDKVKLGKRD 42

## RESULT 4

Q63786 ID Q63786 PRELIMINARY; PRT; 41 AA.  
 AC Q63786;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE Ornithine transcarbamylase (Fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86105223; PubMed=3943133;  
 RA Horwich A.L., Kalousek F., Fenton W.A., Pollock R.A., Rosenberg L.E.;  
 RT Targeting of pre-ornithine transcarbamylase to mitochondria:  
 RT definition of critical regions and residues in the leader peptide.";  
 RL Cell 44:451-459(1986).  
 DR EMBL; M12583; AAA41770.1; -  
 FT NON\_TER 41  
 SQ SEQUENCE 41 AA; 4934 MW; 802465955B32374B CRC64;

Query Match 53.8%; Score 112; DB 11; Length 41;

Best Local Similarity 61.1%; Pred. No. 2.3e-09;

Matches 22; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MLSNRLILLNKAALRKAHTSMVRNFRYKGPVQSOLK 36  
 |||||  
 DB 1 MLFHLRTLLNNAALRNHGFVVRNFRGCGPLQDKVKVQ 36

## RESULT 5

Q9IAU8 ID Q9IAU8 PRELIMINARY; PRT; 354 AA.  
 AC Q9IAU8;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Ornithine transcarbamylase.  
 GN OTC.  
 OS Trachemys scripta elegans.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Testudines; Cryptodira; Testudinidae; Emydidae; Trachemys.  
 OX NCBI\_TaxID=31136;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shimogiri T., Koyanagi K., Mannen H., Tsuji S.;  
 RT "Ornithine transcarbamylase is expressed in uricotelic animals.";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF134846; AAF61410.1; -  
 DR HSSP; P00480; 1OTH.  
 DR InterPro; IPR006130; Asp/Orn\_Cotranf.  
 DR InterPro; IPR002292; Orn\_carbtransf.  
 DR InterPro; IPR006131; OTCace\_O.  
 DR InterPro; IPR006132; OTCace\_P.  
 DR Pfam; PF00185; OTCace; 1.  
 DR Pfam; PF02729; OTCace.N; 1.  
 DR PRINTS; PR00100; AOTCASE.  
 DR TIGRFAMs; TIGR00658; orni\_carb\_tr; 1.  
 DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
 SQ SEQUENCE 354 AA; 39958 MW; BD6A2C3AEC0F99BE CRC64;

Query Match 44.2%; Score 92; DB 13; Length 354;  
 Best Local Similarity 52.4%; Pred. No. 2.6e-05;



DR Pfam; PF00185; OTCace; 1.  
DR Pfam; PF02729; OTCace\_N; 1.  
DR PRINTS; PR00100; AOTCase.  
DR TIGRPFAMS; TIGR00658; Orni\_carb\_tr; 1.  
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
SQ SEQUENCE 356 AA; 40399 MW; 01D49CCB93E4DBD1 CRC64;

Query Match 26.9%; Score 56; DB 13; Length 356;  
Best Local Similarity 38.6%; Pred. No. 6.7;  
Matches 17; Conservative 6; Mismatches 17; Indels 4; Gaps 2;

QY 1 MLSNRLILNKAALRKAHTSMVRNFRY--GKPVQS--OLKPRDL 40  
||| ||| : : : : : ||| : : : ||| |||  
DB 1 MLNFESLFTSRNVNKISHLVQIRYRHGPPSETPVQLGKRL 44  
||| ||| : : : : : ||| : : : ||| |||

RESULT 8  
O23351 PRELIMINARY; PRT; 1286 AA.  
ID ID O23351  
AC AC O23351;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE Hypothetical 143.9 kDa protein.  
GN AT4G14970.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98121113; PubMed=9461215;  
RA Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C.,  
RA Bergkamp R., Dirkse W., van Staveren M., Stiekema W., Drost L.,  
RA Ridley P., Hudson S.A., Patel K., Murphy G., Piffanelli P., Wedler H.,  
RA Wedler J., Wambutt R., Weitzenger T., Pohl T.M., Terry N.,  
RA Glieden J., Villarroel R., De Clerck R., van Montagu M., Lecharny A.,  
RA Auborg S., Gy I., Kreis M., Lao N., Kavanagh T., Hempel S., Kotter P.,  
RA Entian K.D., Rieger M., Schaeffer M., Funk B., Mueller-Auer S.,  
RA Silvey M., James R., Montfort A., Pons A., Puigdomenech P., Douka A.,  
RA Voukelatou E., Milioni D., Hatzopoulos P., Piravandi E., Obermaier B.,  
RA Hilbert H., Duesterhoft A., Moores T., Jones J.D.G., Eneva T.,  
RA Palme K., Benes V., Rechman S., Ansoorge W., Cooke R., Berger C.,  
RA Dalseny M., Voet M., Volckaert G., Newes H.W., Klosterman S.,  
RA Schueller C., Chalwatzis N.  
RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of  
Arabidopsis thaliana."  
FL Nature 391:485-488(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RU EU Arabidopsis sequencing project;  
RL submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z97337; CAB10276.1; -.  
DR EMBL; AL161540; CAB78539.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 1286 AA; 143860 MW; F953B283C53D0DEE CRC64;

Query Match 26.9%; Score 56; DB 10; Length 1286;  
Best Local Similarity 35.3%; Pred. No. 28;  
Matches 12; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 3 SNLRILNKAALRKAHTSMVRNFRYRGKPVQSQLK 36  
||||| : : : : : ||| : : : ||| |||  
DB 107 SNLRMLSSSTTKRDESLVRLNLLVSPQLDIQ 140  
||||| : : : : : ||| : : : ||| |||

RESULT 9  
Q9JUG9 PRELIMINARY; PRT; 138 AA.  
ID ID Q9JUG9  
AC AC Q9JUG9;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

```

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein NMA1316.
GN NMA1316.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=56599;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Davis P., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506(2000).
DR EMBL; AL162755; CAB84568.1; -.
DR InterPro; IPR006522; Tail_comp_S.
DR TIGRFAMs; TIGR01635; tail_comp_S; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 138 AA; 15738 MW; FCOEB880236F2113 CRC64;

Query Match 26.0%; Score 54.5; DB 16; Length 138;
Best Local Similarity 42.9%; Pred. No. 3.9;
Matches 15; Conservative 4; Mismatches 15; Indels 1; Gaps 1;

QY 6 RILLKAALRKHAHTSMVRNFRY-GKPVQSQLKPRD 39
DB 26 RYLLMRRLSETMHTAVKLNFRYAGRPKWLGLKYRD 60

RESULT 10
Q8AVS9
ID Q8AVS9 PRELIMINARY; PRT; 457 AA.
AC Q8AVS9;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Similar to hypothetical protein MGC11993.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC041284; AAH41284.1; -.
KW Hypothetical protein.
SQ SEQUENCE 457 AA; 51021 MW; EE44FB225DD0039B CRC64;

Query Match 26.0%; Score 54; DB 13; Length 457;
Best Local Similarity 46.4%; Pred. No. 18;
Matches 13; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 2 LSNRLILLNKAALRKHAHTSMVRNFRYK 29
DB 135 LWDLRLKNSKACTLHGHTSWKNIEYDK 162

RESULT 11
O96570
ID O96570 PRELIMINARY; PRT; 544 AA.
AC O96570;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)

```

---

```

DE 37CS protein (Protein CS).
GN CG10561 OR ANON-37CS OR CS.
OS Drosophila lebanonensis (Fruit fly) (Scaptodrosophila lebanonensis).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Scaptodrosophila.
OX NCBI_TaxID=7225;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BEIRUT;
RX MEDLINE=99250256; PubMed=10231575;
RA Tataronkov A., Saez A.G., Ayala F.J.;
RT "A compact gene cluster in drosophila: the unrelated Cs gene is
RT compressed between duplicated amd and dac."
RL Gene 231:111-120(1999).
CC -!- FUNCTION: HAS A NONVITAL FUNCTION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
DR EMBL; AF091329; AAC67584.1; -.
DR Flybase; FBgn0025668; Dlebc\CG10561.
DR InterPro; IPR002937; Amino_Oxidase.
DR InterPro; IPR00205; NAD_Binding.
DR Pfam; PF01593; Amino_oxidase; 1.
SQ SEQUENCE 544 AA; 61026 MW; CF003E2CBB7D6DBE CRC64;

Query Match 26.0%; Score 54; DB 5; Length 544;
Best Local Similarity 37.1%; Pred. No. 21;
Matches 13; Conservative 4; Mismatches 18; Indels 0; Gaps 0;

QY 1 MSLNRLILLNKAALRKHAHTSMVRNFRYKGPVQSOL 35
DB 332 VLKNEFALLFPALPLEKLOAIRNLGYGNPKYIL 366

RESULT 12
Q52393
ID Q52393 PRELIMINARY; PRT; 769 AA.
AC Q52393;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE HrpY protein.
GN HRPY.
OS Pseudomonas syringae (pv. phaseolicola).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=319;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Race4;
RX MEDLINE=95178735; PubMed=7873779;
RA Mansfield J., Jenner C., Hockenhull R., Bennett M.A., Stewart R.,
RA Mansfield J., Jenner C., Hockenhull R., Bennett M.A., Stewart R.;
RT "Characterization of avrPphE, a gene for cultivar-specific avirulence
RT from Pseudomonas syringae pv. phaseolicola which is physically linked
RT to hrpY, a new hrp gene identified in the halo-blight bacterium."
RL MOL. Plant Microbe Interact. 7:726-739(1994).
DR EMBL; U16817; AAA67932.1; -.
SQ SEQUENCE 769 AA; 80021 MW; 3BDEB3F41942FBB8 CRC64;

Query Match 26.0%; Score 54; DB 2; Length 769;
Best Local Similarity 36.4%; Pred. No. 32;
Matches 12; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 7 ILLNKAALRKHAHTSMVRNFRYKGPVQSOLKPRD 39
DB 448 LVSSDSALQKAVTEQVQNVNSGKALQTDLATD 480

RESULT 13
O50746
ID O50746 PRELIMINARY; PRT; 260 AA.
AC O50746;
DT 01-JUN-1998 (TReMBLrel. 06, Created)

```

```

DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DI Hypochemical protein BBG23.
GN BBG23.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp28-2.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kierlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer M., Adams M.D., Gocayne J.D., Weidman J.,
RA Uitterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi."
RL Nature 390:580-586(1997).
DR EMBL; AE000786; AAC6607.1; -.
DR TIGR; BBG23; -.
KW Hypochemical protein; Plasmid; Complete proteome.
SQ SEQUENCE 260 AA; 29807 MW; F3F8905AC8829867 CRC64;

Query Match 25.7%; Score 53.5; DB 16; Length 260;
Best Local Similarity 21.8%; Pred. No. 11;
Matches 12; Conservative 14; Mismatches 10; Indels 19; Gaps 2;

QY 6 RLLNKAALRK-----AHTSMVNRFRYKPKVQS-----OLKPRDLC 41
: : : : : : : : : : : : : : : : : : : : : : : : : :
DB 197 RLVTDYSSIRRDNSWSTGTSPTVTHIEYGEIQSWRYSYKYSDDSTIRPKNTC 251
: : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
Q9KAV8 PRELIMINARY; PRT; 383 AA.
AC Q9KAV8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Short-chain-specific acyl-CoA dehydrogenase.
GN BH2178.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirana C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001514; BAB05897.1; -.
DR InterPro; IPR006090; Acyl-CoA_dh_C.
DR InterPro; IPR006091; Acyl-CoA_dh_M.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
KW Complete proteome.
SQ SEQUENCE 383 AA; 42742 MW; E20129A4F567A8C2 CRC64;

Query Match 25.7%; Score 53.5; DB 16; Length 383;
Best Local Similarity 42.9%; Pred. No. 17;
Matches 15; Conservative 3; Mismatches 14; Indels 3; Gaps 1;

QY 2 LSNRLILNKAALRKHTSMVNRFRYKPKVQSOLK 36
: : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

DB 114 IONKOILLNSAATEKATGSPARG---GKPTTTATK 145

RESULT 15
Q8KNN7 PRELIMINARY; PRT; 499 AA.
AC Q8KNN7;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Novyilisin.
GN NVL
OS Clostridium novyi.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1542;
RN [1]
RP SEQUENCE FROM N.A.
RA Cerreto R.R., Valle J.J., Gonzalez-Zorn B.B., Vazquez-Boland J.J.A.;
RT "Clostridium novyi, nvl gene for novyilisin."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ404676; CAC01998.1; -.
DR InterPro; IPR001869; Thiol_cytolysin.
DR Pfam; PF01289; Thiol_cytolysin; 1.
DR PRIN2; PR01400; TACYTOLYSIN.
DR ProDom; PD007062; Thiol_cytolysin; 1.
DR PROSITE; PS00481; THIOI_CYTOLYSINS; 1.
SQ SEQUENCE 499 AA; 55628 MW; 686B917EBC6615F8 CRC64;

Query Match 25.5%; Score 53; DB 2; Length 499;
Best Local Similarity 39.3%; Pred. No. 27;
Matches 11; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 9 LNKALRKHTSMVNRFRYKPKVQSOLK 36
: : : : : : : : : : : : : : : : : : : : : : : : : :
DB 254 LKGVSNAPPLMVSNVAYGRPIYVKLE 281
: : : : : : : : : : : : : : : : : : : : : : : : : :

Search completed: July 24, 2003, 22:07:34
Job time : 34.2857 secs

```

GenCore version 5.1.6  
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 22:02:54 ; Search time 20,9881 seconds  
(without alignments)  
86,686 Million cell updates/sec

Title: us-08-765-244-22

Perfect score: 217

Sequence: 1 MLSNRLILINKAALRKAHTS.....NFRYKPVQSQVLKPRDLC 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	42.4	32	US-08-373-190-29	Sequence 29, Appl
2	92	42.4	32	US-08-438-190A-29	Sequence 29, Appl
3	92	42.4	32	US-08-350-215-29	Sequence 29, Appl
4	92	42.4	32	US-09-287-145A-29	Sequence 29, Appl
5	92	42.4	32	US-09-556-111-29	Sequence 29, Appl
6	54	24.9	2539	US-09-413-814-42	Sequence 42, Appl
7	53	24.4	457	US-09-252-991A-32812	Sequence 32812, A
8	52	24.0	218	US-09-252-991A-18053	Sequence 18053, A
9	51	23.5	1169	US-08-542-921-2	Sequence 2, Appl
10	51	23.5	1169	US-08-880-685-2	Sequence 2, Appl
11	51	23.5	1169	US-08-880-684-2	Sequence 2, Appl
12	49.5	22.8	208	US-09-107-532A-5399	Sequence 5399, Ap
13	49.5	22.8	523	US-09-252-991A-18653	Sequence 18653, A
14	49.5	22.8	1164	US-09-457-708-2	Sequence 2, Appl
15	49.5	22.8	1164	US-09-950-046A-2	Sequence 2, Appl
16	48	22.1	89	US-09-732-210-98	Sequence 98, Appl
17	47	21.7	20	PCT-US95-07543-5	Sequence 5, Appl
18	47	21.7	91	US-09-732-210-195	Sequence 195, App
19	47	21.7	421	US-09-002-567B-1	Sequence 1, Appl
20	47	21.7	421	US-09-002-567B-3	Sequence 3, Appl
21	47	21.7	421	US-09-571-347-1	Sequence 1, Appl
22	47	21.7	421	US-09-571-347-3	Sequence 3, Appl
23	47	21.7	567	US-09-188-811-2	Sequence 2, Appl
24	46.5	21.4	274	US-09-252-991A-29653	Sequence 29653, A
25	46.5	21.4	582	US-09-091-725-17	Sequence 17, Appl
26	46.5	21.4	659	US-09-562-737-16	Sequence 16, Appl
27	46.5	21.4	1711	US-08-365-822C-10	Sequence 10, Appl

Query Match

42.4% Score 92; DB 2; Length 32;

28 46.5 21.4 1711 3 US-08-582-776C-10 Sequence 10, Appl  
29 46.5 21.4 1711 3 US-08-434-831B-10 Sequence 10, Appl  
30 46 21.2 559 1 US-08-313-553-15 Sequence 15, Appl  
31 46 21.2 559 3 US-08-767-993-15 Sequence 15, Appl  
32 46 21.2 729 4 US-09-252-991A-32535 Sequence 32535, A  
33 46 21.2 1139 4 US-09-914-259-68 Sequence 68, Appl  
34 45.5 21.0 761 4 US-09-585-858-27 Sequence 27, Appl  
35 45 20.7 236 4 US-09-252-991A-28901 Sequence 28901, A  
36 45 20.7 275 4 US-09-328-352-4817 Sequence 4817, Ap  
37 45 20.7 535 4 US-09-312-762A-14 Sequence 14, Appl  
38 45 20.7 609 4 US-08-107-532A-5110 Sequence 5110, Ap  
39 45 20.7 700 4 US-08-671-757A-10 Sequence 10, Appl  
40 45 20.7 716 4 US-09-219-983A-20 Sequence 20, Appl  
41 45 20.7 730 1 US-07-846-181-5 Sequence 5, Appl  
42 45 20.7 730 1 US-07-845-989-5 Sequence 5, Appl  
43 45 20.7 795 4 US-09-252-991A-19085 Sequence 19085, A  
44 44 20.7 984 4 US-09-328-352-6926 Sequence 6926, Ap  
45 45 20.7 1124 4 US-09-252-991A-26810 Sequence 26810, A

## ALIGNMENTS

RESULT 1  
US-08-373-190-29  
; Sequence 29, Application US/08373190  
; Patent No. 5851829  
; GENERAL INFORMATION:  
; APPLICANT: MARASCO, WAYNE  
; APPLICANT: HASELTINE, WILLIAM  
; TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF PROTEINS  
; NUMBER OF SEQUENCES: 79  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN  
; STREET: 130 WATER STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/373,190  
; FILING DATE: 17-JAN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/06735  
; FILING DATE: 16-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: RESNICK, DAVID S  
; REGISTRATION NUMBER: 34,235  
; REFERENCE/DOCKET NUMBER: 41956-PCT-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; TELEX: STRE UR 2002  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
US-08-373-190-29

```

CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/350,215
FILING DATE: 12-DEC-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: EISENSTEIN, RONALD I.
REGISTRATION NUMBER: 30628
REFERENCE/DOCKET NUMBER: 41956-CP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 523-3400
TELEFAX: (617) 523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 29:
LENGTH: 32 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-350-215-29

Query Match 42.4% Score 92; DB 3; Length 32;
Best Local Similarity 61.3%; Pred. No. 9.2e-08;
Matches 19; Conservative 2; Mismatches 10; Indels

QY 1 MLSNRLINKAALRKAHTSMVNFRRYKGPV 31
   ||||| ||||| ||||| ||||| |||||
DB 1 MLFNLRXXLLNNAAFRGHGNFVNFNRGQPL 31

RESULT 4
US-09-287-145A-29
Sequence 29, Application US/09287145A
Patent No. 6072036
GENERAL INFORMATION:
APPLICANT: MARASCO, WAYNE
APPLICANT: HASELTINE, WILLIAM
TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
ADDRESSEE: CUSHMAN
STREET: 130 WATER STREET
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,145A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US/08/438,190
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: EISENSTEIN, RONALD I.
REGISTRATION NUMBER: 30628
REFERENCE/DOCKET NUMBER: 41956
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 523-3400

```



RESULT 8  
US-09-252-991A-18053  
; Sequence 18053, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 18053  
; LENGTH: 218  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18053

Query Match 24.0%; Score 52; DB 4; Length 218;  
Best Local Similarity 33.3%; Pred. No. 2.4;  
Matches 16; Conservative 6; Mismatches 20; Indels 6; Gaps 1;  
QY 1 MSLNRLILLNKAALRK-----AHTSMVRNFRYKPVQSQVOLKPRDL 42  
Db 116 MLDLPLKATHAGRRPFISIAHLEVGAEHKGPLASLEQLTPREV 163

RESULT 9  
US-08-542-921-2  
; Sequence 2, Application US/08542921  
; Patent No. 5736514  
; GENERAL INFORMATION:  
; APPLICANT: IIZUKA, TOSHIHIKO  
; APPLICANT: TAGAWA, MICHITO  
; APPLICANT: ARAI, SATOSHI  
; APPLICANT: NIIZEKI, MASATSUGU  
; APPLICANT: MIYAKE, TOSHIRO  
; TITLE OF INVENTION: NOVEL BACILLUS STRAIN AND HARMFUL  
; TITLE OF INVENTION: ORGANISM CONTROLLING AGENTS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/542,921  
; FILING DATE: 13-OCT-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 276082/94  
; FILING DATE: 14-OCT-1994  
; NAME: OBLON, NORMAN F  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 49-209-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1169 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-542-921-2

Query Match 24.0%; Score 52; DB 4; Length 218;  
Best Local Similarity 33.3%; Pred. No. 2.4;  
Matches 16; Conservative 6; Mismatches 20; Indels 6; Gaps 1;  
QY 1 MSLNRLILLNKAALRK-----AHTSMVRNFRYKPVQSQVOLKPRDL 42  
Db 116 MLDLPLKATHAGRRPFISIAHLEVGAEHKGPLASLEQLTPREV 163

INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1169 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-542-921-2  
Query Match 23.5%; Score 51; DB 1; Length 1169;  
Best Local Similarity 25.5%; Pred. No. 28;  
Matches 14; Conservative 12; Mismatches 15; Indels 14; Gaps 2;  
QY 2 LSNRLILLN-----KAALRKAHTSMVRNFRYG---KPVQSQVOLKPRDL 42  
Db 359 LSNLQILVNYQTNSAWRGSRVRYHYLHSSIOEKSYGLLSDPVGANINWQNDI 413

RESULT 10  
US-08-880-685-2  
; Sequence 2, Application US/08880685  
; Patent No. 5834296  
; GENERAL INFORMATION:  
; APPLICANT: IIZUKA, TOSHIHIKO  
; APPLICANT: TAGAWA, MICHITO  
; APPLICANT: ARAI, SATOSHI  
; APPLICANT: NIIZEKI, MASATSUGU  
; APPLICANT: MIYAKE, TOSHIRO  
; TITLE OF INVENTION: NOVEL BACILLUS STRAIN AND HARMFUL  
; TITLE OF INVENTION: ORGANISM CONTROLLING AGENTS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/880,685  
; FILING DATE: 23-JUN-1997  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/542,921  
; FILING DATE: 13-OCT-1995  
; APPLICATION NUMBER: JP 276082/94  
; FILING DATE: 14-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 49-209-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1169 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-880-685-2

Query Match 23.5%; Score 51; DB 2; Length 1169;  
Best Local Similarity 25.5%; Pred. No. 28;  
Matches 14; Conservative 12; Mismatches 15; Indels 14; Gaps 2;  
QY 2 LSNRLILLN-----KAALRKAHTSMVRNFRYG---KPVQSQVOLKPRDL 42  
Db 359 LSNLQILVNYQTNSAWRGSRVRYHYLHSSIOEKSYGLLSDPVGANINWQNDI 413





Halley, Dicky

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 22:01:38 ; Search time 30.7143 seconds  
(without alignments)  
361.274 Million cell updates/sec

Title: US-08-765-244-22

Perfect score: 217

Sequence: 1 MSLNRLILLNKAALKAHTS.....NFRYKGPQSQVQLKPRDL 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_23.\*

- 1: sp-archaea.\*
- 2: sp-bacteria.\*
- 3: sp-fungi.\*
- 4: sp-human.\*
- 5: sp-invertebrate.\*
- 6: sp-mammal.\*
- 7: sp-mhc.\*
- 8: sp-organelle.\*
- 9: sp-phage.\*
- 10: sp-plant.\*
- 11: sp-rodent.\*
- 12: sp-virus.\*
- 13: sp-vertebrate.\*
- 14: sp-unclassified.\*
- 15: sp-virus.\*
- 16: sp-bacteria.\*
- 17: sp-archae.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	199	91.7	102	11	P70518	P70518 rattus norv
2	164	75.6	351	11	Q8RIAB	Q8RIAB mus musculu
3	130	59.9	354	6	Q9N1U7	Q9N1U7 bos taurus
4	119	54.8	41	11	Q63786	Q63786 rattus norv
5	106	48.8	354	13	Q91AU8	Q91AU8 trachemys s
6	70	32.3	356	13	Q91AV0	Q91AV0 sceloporu
7	61	28.1	1286	10	O23351	O23351 arabidopsis
8	60	27.6	351	13	Q97W55	Q97W55 sulfolobus
9	57	26.3	354	17	Q97HY9	Q97HY9 gallus gall
10	56	25.8	369	2	Q92635	Q92635 streptococ
11	55.5	25.6	454	10	Q91MN3	Q91MN3 arabidopsis
12	54	24.9	300	5	O44679	O44679 caenorhabdi
13	54	24.9	457	13	Q8AVS9	Q8AVS9 xenopus lae
14	54	24.9	544	5	O96570	O96570 drosophila
15	54	24.9	866	5	Q9GPG3	Q9GPG3 drosophila
16	54	24.9	866	5	Q9W0E4	Q9W0E4 drosophila

17	54	24.9	1053	5	Q8IRH0	Q8IRH0 drosophila
18	54	24.9	1075	5	Q8IRH1	Q8IRH1 drosophila
19	53.5	24.7	141	16	Q8DG64	Q8DG64 synechococ
20	53.5	24.7	425	2	Q45965	Q45965 coxiella bu
21	53.5	24.7	755	5	Q9V4J5	Q9V4J5 drosophila
22	53.5	24.7	781	5	Q8IGP7	Q8IGP7 drosophila
23	53	24.4	247	10	Q9LQU9	Q9LQU9 oryza sativ
24	53	24.4	593	16	Q8ZC64	Q8ZC64 versinia pe
25	53	24.4	601	16	Q8D0D5	Q8D0D5 versinia pe
26	53	24.4	852	2	Q8KJM4	Q8KJM4 rhizobium l
27	52.5	24.2	140	12	Q99A14	Q99A14 bovine vira
28	52	24.0	213	16	Q9HXS7	Q9HXS7 pseudomonas
29	52	24.0	335	5	O76719	O76719 caenorhabdi
30	52	24.0	357	13	Q9IAU9	Q9IAU9 alligator m
31	52	24.0	598	16	Q8FH22	Q8FH22 escherichia
32	52	24.0	1009	5	Q8WQJ6	Q8WQJ6 drosophila
33	52	24.0	3004	5	Q24550	Q24550 drosophila
34	52	24.0	3004	5	Q9VYN8	Q9VYN8 drosophila
35	51.5	23.7	260	16	O50746	O50746 borrelia bu
36	51	23.5	226	16	Q8EF28	Q8EF28 shewanella
37	51	23.5	376	5	Q8SUE7	Q8SUE7 encephalito
38	51	23.5	499	2	Q8KNM7	Q8KNM7 clostridium
39	51	23.5	598	16	Q8ZP35	Q8ZP35 salmonella
40	51	23.5	598	16	Q8Z7G1	Q8Z7G1 salmonella
41	51	23.5	633	16	P74554	P74554 synechocyst
42	51	23.5	645	2	Q9S603	Q9S603 bacillus th
43	51	23.5	1779	5	O18150	O18150 caenorhabdi
44	50.5	23.3	138	16	Q9JUG9	Q9JUG9 neisseria m
45	50.5	23.3	296	5	Q9VXZ7	Q9VXZ7 drosophila

#### ALIGNMENTS

#### RESULT 1

P70518  
ID P70518 PRELIMINARY; PRT: 102 AA.  
AC P70518;  
DT 01-FEB-1997 (TREMREL. 02, Created)  
DT 01-FEB-1997 (TREMREL. 02, Last sequence update)  
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)  
DE Ornithine carbamoyltransferase precursor (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85051832; PubMed=6548714;  
RA McIntyre P., Graf L., Mercer J., Peterson G., Hudson P.J.,  
RA Hoogenraad N.;  
RT "A highly basic N-terminal extension of the mitochondrial matrix  
enzyme ornithine transcarbamylase from rat liver.";  
RL FEBS Lett. 177:41-46(1984).  
DR EMBL; K03041; AAA41771.1; -;  
DR HSSP; P00480; 10TH.  
DR InterPro; IPR006130; Asp/Orn\_COTranf.  
DR InterPro; IPR006132; Ofcace\_P.  
DR Pfam; PF02729; Ofcace\_N; 1.  
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
KW SIGNAL; Transferase.  
FT SIGNAL 1 32 POTENTIAL.  
FT CHAIN 33 >102 ORNITHINE CARBAMOYLTRANSFERASE.  
FT NON\_TER 102 102  
SQ SEQUENCE 102 AA; B739D41F19BF720D CRC64;

Query Match  
Best Local Similarity 91.7%; Score 199; DB 11; Length 102;  
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MSLNRLILLNKAALKAHTSMVNRFRGKGPVQSQVQLKPRDL 42  
|||||  
DB 1 MSLNRLILLNKAALKAHTSMVNRFRGKGPVQSQVQLKGRDL 42

## RESULT 2

Q8RIA8 ID Q8RIA8 PRELIMINARY; PRT; 351 AA.  
AC Q8RIA8;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Similar to ornithine transcarbamylase.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Strausberg R.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC024893; AAH24893.1; -.  
DR HSSP; P00479; 3CSU.  
DR InterPro; IPR006130; Asp/Orn\_Cotranf.  
DR InterPro; IPR002292; Orn\_carbtransf.  
DR InterPro; IPR006131; OTCace\_O.  
DR InterPro; IPR006132; OTCace\_P.  
DR Pfam; PF00185; OTCace; 1.  
DR Pfam; PF02729; OTCace.N; 1.  
DR TIGRFAMS; TIGR00658; orni\_carb.tr; 1.  
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
SQ SEQUENCE 351 AA; 39365 MW; E6B38B2FC779F42 CRC64;

Query Match 75.6%; Score 164; DB 11; Length 351;  
Best Local Similarity 83.3%; Pred. No. 5.9e-16;  
Matches 35; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MLSNLRILLNKAALRKAHTSMVNRFRYKPKVQSQVQLKPRDL 42  
|||||  
DB 1 MLSNLRILLNKAALRKAHTSVVRFHFCWCKPKVQSQVQLKGRDL 42

## RESULT 3

Q9NIU7 ID Q9NIU7 PRELIMINARY; PRT; 354 AA.  
AC Q9NIU7;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Ornithine transcarbamylase.  
GN OTC.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Shimogiri T., Koyanagi K., Mannen H., Tsuji S.;  
RT "Ornithine transcarbamylase is expressed in uricotelic animals.";  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF134841; AAF61405.1; -.  
DR HSSP; P00480; 1OTH.  
DR InterPro; IPR006130; Asp/Orn\_Cotranf.  
DR InterPro; IPR002292; Orn\_carbtransf.  
DR InterPro; IPR006131; OTCace\_O.  
DR InterPro; IPR006132; OTCace\_P.  
DR Pfam; PF00185; OTCace; 1.  
DR Pfam; PF02729; OTCace.N; 1.  
DR PRINTS; PR00100; AOTCASE.  
DR TIGRFAMS; TIGR00658; orni\_carb.tr; 1.  
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
SQ SEQUENCE 354 AA; 39842 MW; 30EAC2EF4AC1D71B CRC64;

Query Match 59.9%; Score 130; DB 6; Length 354;

Best Local Similarity 64.3%; Pred. No. 7.1e-11;  
Matches 27; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNLRILLNKAALRKAHTSMVNRFRYKPKVQSQVQLKPRDL 42  
|||||  
DB 1 MLFHLRTLNNAAALRNHFNWFRFCGQPLQDKVQLKGRDL 42

## RESULT 4

Q63786 ID Q63786 PRELIMINARY; PRT; 41 AA.  
AC Q63786;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE Ornithine transcarbamylase (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86106223; Pubmed=3943133;  
RA Horwich A.L., Kalousek F., Fenton W.A., Pollock R.A., Rosenberg L.E.;  
RT Targeting of pre-ornithine transcarbamylase to mitochondria:  
RT definition of critical regions and residues in the leader peptide.";  
RL Cell 44:451-459(1986).  
DR EMBL; M12583; AAA41770.1; -.  
DR NON\_TER 41  
FT NON\_TER 41  
SQ SEQUENCE 41 AA; 4934 MW; 802465955B32374B CRC64;

Query Match 54.8%; Score 119; DB 11; Length 41;  
Best Local Similarity 66.7%; Pred. No. 2.8e-10;  
Matches 24; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 MLSNLRILLNKAALRKAHTSMVNRFRYKPKVQSQVQ 36  
|||||  
DB 1 MLFNLRLNNAAAFRNHFNWFRFCGQPLQNKVQ 36

## RESULT 5

Q9IAU8 ID Q9IAU8 PRELIMINARY; PRT; 354 AA.  
AC Q9IAU8;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Ornithine transcarbamylase.  
GN OTC.  
OS Trachemys scripta elegans.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Testudines; Cryptodira; Testudinoidae; Emydidae; Trachemys.  
OX NCBI\_TaxID=31138;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Shimogiri T., Koyanagi K., Mannen H., Tsuji S.;  
RT "Ornithine transcarbamylase is expressed in uricotelic animals.";  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF134846; AAF61410.1; -.  
DR HSSP; P00480; 1OTH.  
DR InterPro; IPR006130; Asp/Orn\_Cotranf.  
DR InterPro; IPR002292; Orn\_carbtransf.  
DR InterPro; IPR006131; OTCace\_O.  
DR InterPro; IPR006132; OTCace\_P.  
DR Pfam; PF00185; OTCace; 1.  
DR Pfam; PF02729; OTCace.N; 1.  
DR PRINTS; PR00100; AOTCASE.  
DR TIGRFAMS; TIGR00658; orni\_carb.tr; 1.  
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
SQ SEQUENCE 354 AA; 39958 MW; BD6AZC3AEC0F99BE CRC64;

Query Match 48.8%; Score 106; DB 13; Length 354;  
Best Local Similarity 52.4%; Pred. No. 2.7e-07;

Matches 22; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 MLSNRLINKAALKAHSMVNFYRGKPVQSQVQLKPRDL 42  
 DB 1 MLFNRLNLAATLRNSSQQLVQHFSGQPTTNINLKGRL 42

RESULT 6  
 Q9IAVO SEQUENCE FROM N.A.  
 ID Q9IAVO PRELIMINARY; PRT; 356 AA.  
 AC Q9IAVO;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Ornithine transcarbamylase.  
 GN OTC.  
 OS Sceloporus undulatus (Eastern fence lizard) (Stink).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Iguania; Iguanidae; Phrynosomatinae;  
 OC Sceloporus.  
 OX NCBI\_TaxID=8520;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shimogiri T., Koyanagi K., Mannen H., Tsuji S.;  
 RT "Ornithine transcarbamylase is expressed in uricotelic animals.";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF134844; AAF61408.1; -;  
 DR HSSP; P00480; IOTH.  
 DR InterPro; IPR006130; Asp/Oro\_Cotransf.  
 DR InterPro; IPR002292; Orn\_carbtransf.  
 DR InterPro; IPR006131; OTCace.O.  
 DR InterPro; IPR006132; OTCace.P.  
 DR Pfam; PF00185; OTCace.1.  
 DR Pfam; PF02729; OTCace\_N; 1.  
 DR PRINTS; PR00100; AOTCASE.  
 DR TIGRFAMs; TIGR00658; Orn\_carb\_tr; 1.  
 DR PROSITE; PS00097; CARBAMYLTRANSFASRASE; 1.  
 SQ SEQUENCE 356 AA; 40399 MW; 01D49CCB93E4BBD1 CRC64;

Query Match 32.3%; Score 70; DB 13; Length 356;  
 Best Local Similarity 40.9%; Pred. No. 0.063;  
 Matches 18; Conservative 6; Mismatches 18; Indels 2; Gaps 1;

QY 1 MLSNRLINKAALKAHSMVNFYRGKPVQSQVQLKPRDL 42  
 DB 1 MLFNRLNLAATLRNSSQQLVQHFSGQPTTNINLKGRL 42

RESULT 7  
 O23351 SEQUENCE FROM N.A.  
 ID O23351 PRELIMINARY; PRT; 1286 AA.  
 AC O23351;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE Hypothetical 143.9 kDa protein.  
 GN AT4G14970.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C.,  
 RA Bergkamp R., Dirkes W., van Staveren M., Stiekema W., Drost L.,  
 RA Ridley P., Hudson S.A., Patel K., Murphy G., Piffanelli P., Wedler H.,  
 RA Wedler E., Wambutt R., Weitzengraber T., Pohl T.M., Terry N.,  
 RA Gieren J., Villarroel R., De Clerck R., van Montagu M., Lecharny A.,  
 RA Auborg S., Gy I., Kreis M., Lao N., Kavanagh T., Hempel S., Kotter P.,  
 RA Entian K.D., Rieger M., Schaeffer M., Funk B., Mueller-Auer S.,  
 RA Silvey M., James R., Montfort A., Pons A., Puigdomenech P., Douka A.,

RA Voukelatou E., Milioni D., Hatzopoulos P., Piravandi E., Obermaier B.,  
 RA Hilbert H., Duesterhoft A., Moors A., Jones J.D.G., Eneva T.,  
 RA Paine K., Benes V., Rechman S., Ansoorge W., Cooke R., Berger C.,  
 RA Delseny M., Voet M., Volckaert G., Mewes H.W., Klosterman S.,  
 RA Schueller C., Chalvatizis N.;  
 RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of  
 Arabidopsis thaliana.";  
 RL Nature 391:485-488(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z97337; CAB10276.1; -;  
 DR EMBL; AL161540; CAB78539.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 1286 AA; 143860 MW; F953B283C53D0DEE CRC64;

Query Match 28.1%; Score 61; DB 10; Length 1286;  
 Best Local Similarity 38.2%; Pred. No. 5.8;  
 Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 3 SNLRLLINKAALKAHSMVNFYRGKPVQSQV 36  
 DB 107 SNLRMLSSSTTKRDESLVRLNLLVSPQLDIO 140

RESULT 8  
 Q97W55 SEQUENCE FROM N.A.  
 ID Q97W55 PRELIMINARY; PRT; 351 AA.  
 AC Q97W55;  
 DT 01-OCT-2001 (TREMBLrel. 18, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Hypothetical GTP binding protein SSO2385.  
 GN SSO2385.  
 OS Sulfolobus solfataricus.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC Sulfolobus.  
 OX NCBI\_TaxID=2287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE=21332296; PubMed=11427726;  
 RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,  
 RA Awey M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,  
 RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,  
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,  
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,  
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,  
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).  
 DR EMBL; AE006839; AAK42533.1; -;  
 DR InterPro; IPR004095; TGS\_dom.  
 DR Pfam; PF02824; TGS; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 351 AA; 39916 MW; 046A96BF004865DE CRC64;

Query Match 27.6%; Score 60; DB 17; Length 351;  
 Best Local Similarity 34.6%; Pred. No. 1.9;  
 Matches 18; Conservative 8; Mismatches 10; Indels 16; Gaps 3;

QY 7 ILLNKA-----LRKAHSMVNFYRGKPVQSQV-----LKPRDL 42  
 DB 294 LILKAGTVLDVARKLHSSLAENFYRVWVGSKVFGQKVGPSHILEDRI 345

RESULT 9  
 Q9YHY9 SEQUENCE FROM N.A.  
 ID Q9YHY9 PRELIMINARY; PRT; 354 AA.  
 AC Q9YHY9;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Ornithine transcarbamylase precursor (EC 2.1.3.3).  
 GN OTC.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BREED WHITE LEHORN; TISSUE=Kidney;  
 RX MEDLINE=99011321; PubMed=9792920;  
 RA Shimogiri T., Kono M., Mannen H., Mizutani M., Tsuji S.;  
 RT "Chicken ornithine transcarbamylase gene, structure, regulation, and  
 RT chromosomal assignment: repetitive sequence motif in intron 3  
 RT regulates this enzyme activity.";  
 RL J. Biochem. 124:962-971(1998).  
 DR EMBL; AF065629; AAD12234.1; .  
 DR EMBL; AF065638; AAD33083.1; .  
 DR EMBL; AF065630; AAD33083.1; JOINED.  
 DR EMBL; AF065631; AAD33083.1; JOINED.  
 DR EMBL; AF065632; AAD33083.1; JOINED.  
 DR EMBL; AF065634; AAD33083.1; JOINED.  
 DR EMBL; AF065635; AAD33083.1; JOINED.  
 DR EMBL; AF065636; AAD33083.1; JOINED.  
 DR EMBL; AF065637; AAD33083.1; JOINED.  
 DR HSSP; P00480; 1OTH.  
 DR InterPro; IPR006130; Asp/Orn\_COTranf.  
 DR InterPro; IPR002299; Orn\_Carboxyltransf.  
 DR InterPro; IPR006131; OTCace.O.  
 DR InterPro; IPR006132; OTCace.P.  
 DR Pfam; PF00185; OTCace; 1.  
 DR Pfam; PF02729; OTCace.N; 1.  
 DR PRINTS; PR00100; AOTCASE.  
 DR TIGRGRAMS; TIGR00658; orni\_carb.tr; 1.  
 DR PROSITE; PS00037; CARBAMOYLTRANSFERASE; 1.  
 DR TRANSFERASE.  
 KW Transferase.  
 SQ SEQUENCE 354 AA; 40245 MW; 20447180BAD9D4ED CRC64;  
  
 Query Match 26.3%; Score 57; DB 13; Length 354;  
 Best Local Similarity 40.5%; Pred. No. 5.5;  
 Matches 17; Conservative 4; Mismatches 21; Indels 0; Gaps 0;  
  
 QY 1 MSLNRLINKAALRKHAHTSMVRNFRYKPVQSQVOLKPRDL 42  
 DB 1 MLFNKLNRYITKLTNSKHLPRHFCRGPNNVCLGRDL 42  
  
 RESULT 10  
 Q92635  
 ID Q92635 PRELIMINARY; PRT; 369 AA.  
 AC Q92635;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Putative glucosyl transferase.  
 GN CPS19CS.  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99395043; PubMed=10464207;  
 RA Morona J.K., Morona R., Paton J.C.;  
 RT "Comparative genetics of capsular polysaccharide biosynthesis in  
 RT Streptococcus pneumoniae types belonging to serogroup 19.";  
 RL J. Bacteriol. 181:5355-5364(1999).  
 DR EMBL; AF105116; AAD1925.1; .  
 DR InterPro; IPR001296; Glyco.trans.1.  
 DR Pfam; PF00534; Glycos\_transf.1; 1.  
 DR TRANSFERASE.  
 KW Transferase.  
 SQ SEQUENCE 369 AA; 43229 MW; FCFDFD5C106AC8BE CRC64;

Query Match 25.8%; Score 56; DB 2; Length 369;  
 Best Local Similarity 32.5%; Pred. No. 8.1;  
 Matches 13; Conservative 9; Mismatches 14; Indels 4; Gaps 1;

QY 4 NLRILINKAALRKHAHTSMVRNFRYKPVQSQVOLKPRDL 43  
 DB 48 NVHKVLVRLGIKKSDMSMT---YIKYAEHQVHLSPEVC 83

RESULT 11  
 Q9LMN3  
 ID Q9LMN3 PRELIMINARY; PRT; 454 AA.  
 AC Q9LMN3;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE F1694.13 protein.  
 GN F1694.13.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Liu S.X., Sakano H., Yu G., Etgu P., Lee J.M., Lenz C., Pham P.,  
 RA Toriumi M., Byun E., Chan A., Chin C., Choi E., Chung M., Altafi H.,  
 RA Goldsmith A., Gonzalez A., Liu A., Smith A., Vaysberg M., Altafi H.,  
 RA Brooks S., Buehler E., Chao Q., Conn L., Conway A.B., Hansen N.F.,  
 RA Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M., Nguyen M.,  
 RA Palm C.J., Shinn P., Southwick A., Davis R.W., Ecker J.R.,  
 RA Federspiel N.A., Theologis A.;  
 RT "The sequence of BAC F1694 from Arabidopsis thaliana chromosome 1.";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC036104; AAF81364.1; .  
 SQ SEQUENCE 454 AA; 50928 MW; 70BC2ACCC9DBEDCF CRC64;

Query Match 25.6%; Score 55.5; DB 10; Length 454;  
 Best Local Similarity 45.2%; Pred. No. 12;  
 Matches 14; Conservative 5; Mismatches 11; Indels 1; Gaps 1;

QY 11 KAALRKHAHT-SMVRNFRYKPVQSQVOLKPR 40  
 DB 422 KAAYSRTSDSPSRNYRESQPMGSPVQARPR 452

RESULT 12  
 O44679  
 ID O44679 PRELIMINARY; PRT; 300 AA.  
 AC O44679;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical 33.8 kDa protein.  
 GN C14C6.13.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
 OC Rhabditidae; Peioderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=Bristol N2;  
 RC MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA David M., Wohlmann P., Bauer C., Antoniou B.;



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: July 24, 2003, 21:59:48 ; Search time 7.67857 Seconds  
(without alignments)  
263.350 Million cell updates/sec

Title: US-08-765-244-22

Perfect score: 217

Sequence: 1 MLNLRLLKKAALRAKHTS.....NFRYKPVQSQVQLKPRDLIC 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	199	91.7	354	1	P00481 rattus norv
2	164	75.6	354	1	P11725 mus musculu
3	141	65.0	354	1	P00480 homo sapien
4	59	27.2	597	1	IF7P_METAC
5	56	25.8	591	1	IF2P_METMA
6	53	25.3	350	1	OTC_RANCA
7	54	24.9	236	1	RR2_LOTJA
8	54	24.9	457	1	HENN_SALTY
9	54	24.9	1225	1	KFAA_CHICK
10	53	24.4	328	1	OTC_PIG
11	53	24.4	435	1	CG55_YEAST
12	52.5	24.2	185	1	EF7L_CHLMU
13	52	24.0	598	1	NARX_ECOLI
14	51	23.5	506	1	POFB_SCHPO
15	51	23.5	705	1	YM37_YEAST
16	51	23.5	1169	1	C9DA_BACTP
17	50.5	23.3	429	1	TRE2_SULSO
18	50	23.0	330	1	LARU_VIBPA
19	50	23.0	622	1	SR68_CANFA
20	50	23.0	662	1	PRCA_ANASP
21	49.5	22.8	864	1	RA50_SULSO
22	49.5	22.8	1163	1	TSC1_HUMAN
23	49.5	22.8	1164	1	TSC1_HUMAN
24	49	22.6	457	1	HENN_ECOLI
25	49	22.6	1219	1	MY55_YEAST
26	49	22.6	1345	1	YH00_YEAST
27	48.5	22.4	5035	1	RYR1_PIG
28	48.5	22.4	5037	1	RYR1_RABIT
29	48.5	22.4	5038	1	RYR1_HUMAN
30	48	22.1	88	1	R37A_SCHPO
31	48	22.1	578	1	SYR_BUCBP
32	48	22.1	647	1	PRCA_ANAVA
33	48	22.1	1025	1	BGAL_KLJLA

## ALIGNMENTS

### RESULT 1

ID	OTC_RAT	STANDARD;	PRT;	354 AA.
AC	P00481; Q63407;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)			
DE	(Ornithine transcarbamylase).			
GN	OTC.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=85063800; PubMed=6095294;			
RA	Takiguchi M., Miura S., Mori M., Tatibana M., Nagata S., Kaziro Y.;			
RT	"Molecular cloning and nucleotide sequence of cDNA for rat ornithine carbamoyltransferase precursor."			
RL	Proc. Natl. Acad. Sci. U.S.A. 81:7412-7416(1984).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Wistar; TISSUE=Liver;			
EX	MEDLINE=87317609; PubMed=3476935;			
RA	Takiguchi M., Murakami T., Miura S., Mori M.;			
RT	"Structure of the rat ornithine carbamoyltransferase gene, a large, x chromosome-linked gene with an atypical promoter."			
RL	Proc. Natl. Acad. Sci. U.S.A. 84:6136-6140(1987).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=85215524; PubMed=3839075;			
RA	Kraus J.P., Hodges P.E., Williamson C.L., Horwich A.L., Kalousek F.,			
RT	Williams K.R., Rosenberg L.E.;			
RT	"A cDNA clone for the precursor of rat mitochondrial ornithine transcarbamylase: comparison of rat and human leader sequences and conservation of catalytic sites."			
RL	Nucleic Acids Res. 13:943-952(1985).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=85203360; PubMed=3838931;			
RA	McIntyre P., Graf L., Mercer J.F.B., Wake S.A., Hudson P.J.,			
RT	Hoogenraad N.;			
RT	"The primary structure of the imported mitochondrial protein, ornithine transcarbamylase from rat liver: mRNA levels during ontogeny."			
RL	DNA 4:147-156(1985).			
RN	[5]			
RP	SEQUENCE OF 33-56; 293-302; 307-317 AND 321-329.			
RC	TISSUE=Liver;			
EX	MEDLINE=88266748; PubMed=3390141;			
RA	Aoki Y., Sunaga H., Suzuki K.T.;			
RT	"A cadmium-binding protein in rat liver identified as ornithine carbamoyltransferase."			
RL	Biochem. J. 250:735-742(1988).			
CC	-!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine - phosphate			

34	47.5	21.9	333	1	Y376_BUCAP	Q8K9g1 buchnera ap
35	47.5	21.9	1473	1	OVOS_CHICK	P20740 gallus gall
36	47.5	21.9	2210	1	RRPL_EBOSM	Q66802 ebola virus
37	47	21.7	90	1	R37B_SCHPO	P05733 schizosacch
38	47	21.7	260	1	PSMA_PYPAB	Q9v122 pyrococcus
39	47	21.7	260	1	PSMA_PYRHO	O59219 pyrococcus
40	47	21.7	421	1	TES_HUMAN	Q9ug18 homo sapien
41	47	21.7	423	1	TES_MOUSE	P47226 mus musculu
42	47	21.7	1098	1	RPOP_MAIZE	P10581 zea mays (m
43	46.5	21.4	410	1	T801_PSESH	P24607 pseudomonas
44	46.5	21.4	421	1	VCOM_ADECC	Q65952 canine aden
45	46.5	21.4	421	1	VCOM_ADECR	Q96685 canine aden

```

CC CC      + L-citrulline.
CC CC      -|- PATHWAY: SECOND STEP IN UREA CYCLE, ARGININE BIOSYNTHESIS.
CC CC      -|- SUBUNIT: Homotrimer.
CC CC      -|- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC CC      -|- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
CC CC      -----
CC CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC      the European Bioinformatics Institute. There are no restrictions on its
CC CC      use by non-profit institutions as long as its content is in no way
CC CC      modified and this statement is not removed. Usage by and for commercial
CC CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC      or send an email to license@isb-sib.ch).
CC CC      -----
CC CC      EMBL: K03040; AAA41768.1; -
CC CC      EMBL: M16933; AAA41769.1; -
CC CC      EMBL: M16924; AAA41769.1; JOINED.
CC CC      EMBL: M16925; AAA41769.1; JOINED.
CC CC      EMBL: M16926; AAA41769.1; JOINED.
CC CC      EMBL: M16928; AAA41769.1; JOINED.
CC CC      EMBL: M16929; AAA41769.1; JOINED.
CC CC      EMBL: M16930; AAA41769.1; JOINED.
CC CC      EMBL: M16932; AAA41769.1; JOINED.
CC CC      EMBL: X01976; CAA26007.1; -
CC CC      EMBL: K00001; AAA41772.1; -
CC CC      EMBL: M11266; AAA41767.1; -
CC CC      PIR: A00563; OWRT.
CC CC      HSSP: P00480; 10TH.
CC CC      InterPro: IPR006130; Asp/Orn_Cotranf.
CC CC      InterPro: IPR002292; Orn_carbtransf.
CC CC      InterPro: IPR006131; OTCace_O.
CC CC      InterPro: IPR006132; OTCace_P.
CC CC      Pfam: PF00185; OTCace; 1.
CC CC      Pfam: PF02729; OTCace_N; 1.
CC CC      PRINTS: PR00100; AOTCASE.
CC CC      TIGRFS: TIGR00658; orni_carb.tr; 1.
CC CC      PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.
CC CC      Arginine biosynthesis; Urea cycle; Transferase; Mitochondrion;
CC CC      Arginine biosynthesis; Urea cycle; Transferase; Mitochondrion;
CC CC      TRANSIT peptide.
CC CC      CHAIN 33 354 MITOCHONDRION.
CC CC      ACT_SITE 263 263 ORNITHINE CARBAMOYLTRANSFERASE.
CC CC      BY SIMILARITY.
CC CC      ACT_SITE 303 303 BY SIMILARITY.
CC CC      G -> P (IN REF. 3); AAA41772).
CC CC      G -> S (IN REF. 3).
CC CC      CONFLICT 241 241
CC CC      SEQUENCE 354 AA; 39886 MW; 156B511AF7063F0C CRC64;

Query Match          91.7%; Score 199; DB 1; Length 354;
Best Local Similarity 97.6%; Pred. No. 9,2e-21;
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MLSNRLILKALRKHAHTSMVNFYGRKPVQSQVOLKPRDL 42
Dd 1 MLSNRLILKALRKHAHTSMVNFYGRKPVQSQVOLKGRDL 42

RESULT 2
OTC_MOUSE
ID OTC_MOUSE STANDARD; PRT; 354 AA.
AC P11725;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)
DE (OTCase) (Ornithine transcarbamylase).
GN OTC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;

MEDLINE=87263407; PubMed=3603027;
Veres G., Gibbs R.A., Scherer S.E., Caskey C.T.;
"The molecular basis of the sparse fur mouse mutation.";
Science 237:415-417(1987).
[2]
SEQUENCE FROM N.A.
MEDLINE=88157717; PubMed=2831503;
Scherer S.E., Veres G., Caskey C.T.;
"The genetic structure of mouse ornithine transcarbamylase.";
Nucleic Acids Res. 16:1593-1601(1988).
[3]
SEQUENCE OF 1-26 FROM N.A.
STRAIN=C57BL/6J;
MEDLINE=86224037; PubMed=3011788;
Veres G., Craigen W.J., Caskey C.T.;
"The 5' flanking region of the ornithine transcarbamylase gene
contains DNA sequences regulating tissue-specific expression.";
J. Biol. Chem. 261:7588-7591(1986).
-|- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
+ L-citrulline.
-|- PATHWAY: SECOND STEP IN UREA CYCLE, ARGININE BIOSYNTHESIS.
-|- SUBUNIT: Homotrimer.
-|- SUBCELLULAR LOCATION: Mitochondrial matrix.
-|- DISEASE: SPARSE FUR (SPF) MOUSE HAVE AN OTC WITH AN OVERALL
DECREASE IN ACTIVITY, AND ALTERED SUBSTRATE AFFINITY.
-|- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: M17030; AAA39865.1; -
EMBL: M12716; AAA39864.1; ALT_SEQ.
EMBL: X07092; CAA30121.1; -
EMBL: X07093; CAA30121.1; JOINED.
EMBL: X07094; CAA30121.1; JOINED.
EMBL: X07095; CAA30121.1; JOINED.
EMBL: X07096; CAA30121.1; JOINED.
EMBL: X07097; CAA30121.1; JOINED.
EMBL: X07098; CAA30121.1; JOINED.
EMBL: X07099; CAA30121.1; JOINED.
EMBL: X07100; CAA30121.1; JOINED.
PIR: A43609; OWMS.
HSSP: P00480; 10TH.
SWISS-2DPAGE: P11725; MOUSE.
MGD: MGI:97448; Otc.
InterPro: IPR006130; Asp/Orn_Cotranf.
InterPro: IPR002292; Orn_carbtransf.
InterPro: IPR006131; OTCace_O.
InterPro: IPR006132; OTCace_P.
Pfam: PF00185; OTCace; 1.
Pfam: PF02729; OTCace_N; 1.
PRINTS: PR00100; AOTCASE.
TIGRFS: TIGR00658; orni_carb.tr; 1.
PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.
Arginine biosynthesis; Urea cycle; Transferase; Mitochondrion;
Transit peptide.
CHAIN 33 354 ORNITHINE CARBAMOYLTRANSFERASE.
BY SIMILARITY.
ACT_SITE 263 263 BY SIMILARITY.
ACT_SITE 303 303 BY SIMILARITY.
H -> N (IN SPARSE FUR MOUSE).
G -> R (IN REF. 2).
CONFLICT 195 195
SEQUENCE 354 AA; 39765 MW; 33BB5D1E8AA196 CRC64;

Query Match          75.6%; Score 164; DB 1; Length 354;
Best Local Similarity 83.3%; Pred. No. 8,8e-16;
Matches 35; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```



QY 1 MSLNRLILLNKAALRKAHTSVNRFRYKGPVQSOVLKPRDL 42  
 Db 1 MSLNRLILLNNAALRKGHTSVRFHWCCKPQSOVLKGRDL 42

## RESULT 3

OTC\_HUMAN OTC\_HUMAN STANDARD; PRT; 354 AA.  
 AC P00480;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)  
 DE (OrnCase) (Ornithine transcarbamylase).  
 GN OTC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=98227905; PubMed=2836378;  
 RA Hata A., Tsuzuki T., Shimada K., Takiguchi M., Mori M., Matsuda I.;  
 RT "Structure of the human ornithine transcarbamylase gene.";  
 RL J. Biochem. 103:302-308(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=84196410; PubMed=6372096;  
 RA Horwich A.L., Fenton W.A., Williams K.R., Kalousek F., Kraus J.P.,  
 RA Doolittle R.F., Konigsberg W., Rosenbery L.E.;  
 RT "Structure and expression of a complementary DNA for the nuclear  
 RT coded precursor of human mitochondrial ornithine transcarbamylase.";  
 RL Science 224:1068-1074(1984).  
 RN [3]  
 RP SEQUENCE OF 1-36 FROM N.A.  
 RX MEDLINE=85270440; PubMed=3895227;  
 RA Horwich A.L., Kalousek F., Rosenbery L.E.;  
 RT "Arginine in the leader peptide is required for both import and  
 RT proteolytic cleavage of a mitochondrial precursor.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4930-4933(1985).  
 RN [4]  
 RP SEQUENCE OF 1-26 FROM N.A.  
 RX MEDLINE=87057134; PubMed=3782067;  
 RA Hata A., Tsuzuki T., Shimada K., Takiguchi M., Mori M., Matsuda I.;  
 RT "Isolation and characterization of the human ornithine  
 RT transcarbamylase gene: structure of the 5'-end region.";  
 RL J. Biochem. 100:717-725(1986).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).  
 RX MEDLINE=99069419; PubMed=9852088;  
 RA Shi D., Morizono H., Ha Y., Aoyagi M., Tuchman M., Allewell N.M.;  
 RT "1.85-A resolution crystal structure of human ornithine  
 RT transcarbamylase complexed with N-phosphonacetyl-L-ornithine.  
 RT Catalytic mechanism and correlation with inherited deficiency.";  
 RL J. Biol. Chem. 273:34247-34254(1998).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=20274073; PubMed=10813810;  
 RA Shi D., Morizono H., Aoyagi M., Tuchman M., Allewell N.M.;  
 RT "Crystal structure of human ornithine transcarbamylase complexed with  
 RT carbamoyl phosphate and L-norvaline at 1.9 A resolution.";  
 RL Proteins 39:271-277(2000).  
 RN [7]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=93372868; PubMed=8364586;  
 RA Tuchman M.;  
 RT "Mutations and polymorphisms in the human ornithine transcarbamylase  
 RT gene.";  
 RL Hum. Mutat. 2:174-178(1993).  
 RN [8]  
 RP REVIEW ON VARIANTS.

RX MEDLINE=95353279; PubMed=7627182;  
 RA Tuchman M., Plante R.J.;  
 RT "Mutations and polymorphisms in the human ornithine transcarbamylase  
 RT gene: mutation update addendum.";  
 RL Hum. Mutat. 5:293-295(1995).  
 RN [9]  
 RP REVIEW ON VARIANTS, AND 3D-STRUCTURE MODELING.  
 RX MEDLINE=96091868; PubMed=8544185;  
 RA Tuchman M., Morizono H., Reish O., Yuan X., Allewell N.M.;  
 RT "The molecular basis of ornithine transcarbamylase deficiency:  
 RT modelling the human enzyme and the effects of mutations.";  
 RL J. Med. Genet. 32:680-688(1995).  
 RN [10]  
 RP VARIANT OTCD GLN-141.  
 RX MEDLINE=89008892; PubMed=3170748;  
 RA Maddalena A., Spence J.E., O'Brien W.E., Nusbaum R.L.;  
 RT "Characterization of point mutations in the same arginine codon in  
 RT three unrelated patients with ornithine transcarbamylase  
 RT deficiency.";  
 RL J. Clin. Invest. 82:1353-1358(1988).  
 RN [11]  
 RP VARIANTS OTCD GLN-26; PRO-45 AND GLU-216, AND VARIANT ARG-46.  
 RX MEDLINE=89345570; PubMed=2474822;  
 RA Grompe M., Muzny D.M., Caskey C.T.;  
 RT "Scanning detection of mutations in human ornithine transcarbamylase  
 RT by chemical mismatch cleavage.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:5888-5892(1989).  
 RN [12]  
 RP VARIANT OTCD TRP-277.  
 RX MEDLINE=90269805; PubMed=2347583;  
 RA Finkelstein J.E., Francmano C.A., Brusilow S.W., Travstman M.D.;  
 RT "Use of denaturing gradient gel electrophoresis for detection of  
 RT mutation and prospective diagnosis in late onset ornithine  
 RT transcarbamylase deficiency.";  
 RL Genomics 7:167-172(1990).  
 RN [13]  
 RP VARIANTS OTCD GLN-92 AND LEU-320, AND VARIANT PRO-111.  
 RX MEDLINE=91118929; PubMed=1671317;  
 RA Grompe M., Caskey C.T., Fenwick R.G. Jr.;  
 RT "Improved molecular diagnostics for ornithine transcarbamylase  
 RT deficiency.";  
 RL Am. J. Hum. Genet. 48:212-222(1991).  
 RN [14]  
 RP VARIANT OTCD LEU-225.  
 RX MEDLINE=92098086; PubMed=1721894;  
 RA Hentzen D., Pelet A., Feldman D., Rabier D., Berthelot J.,  
 RA Munnich A.;  
 RT "Fatal hyperammonemia resulting from a C-to-T mutation at a MspI site  
 RT of the ornithine transcarbamylase gene.";  
 RL Hum. Genet. 88:153-156(1991).  
 RN [15]  
 RP VARIANTS OTCD GLU-79; THR-94; PHE-304 AND ASP-345.  
 RX MEDLINE=93126062; PubMed=1480464;  
 RA Tuchman M., Holzknecht R.A., Gueron A.B., Berry S.A., Tsai M.Y.;  
 RT "Six new mutations in the ornithine transcarbamylase gene detected by  
 RT single-strand conformational polymorphism.";  
 RL Pediatr. Res. 32:600-604(1992).  
 RN [16]  
 RP SEQUENCE OF 269-289 FROM N.A., AND VARIANT OTCD GLN-277.  
 RX MEDLINE=94362689; PubMed=8081373;  
 RA Gilbert-Dussardier B., Rabier D., Strautnieks S., Segues B.,  
 RA Bonnefont J.-P., Munnich A.;  
 RT "A novel arginine (245) to glutamine change in exon 8 of the ornithine  
 RT carbamoyl transferase gene in two unrelated children presenting with  
 RT late onset deficiency and showing the same enzymatic pattern.";  
 RL Hum. Mol. Genet. 3:831-832(1994).  
 RN [17]  
 RP VARIANT OTCD PRO-140.  
 RX MEDLINE=93373296; PubMed=8099056;  
 RA Tsai M.Y., Holzknecht R.A., Tuchman M.;  
 RT "Single-strand conformational polymorphism and direct sequencing  
 RT applied to carrier testing in families with ornithine  
 RT transcarbamylase deficiency.";

```

RL Hum. Genet. 91:321-325(1993).
RN [18]
RP VARIANTS OTCD LEU-117; LEU-182 AND CYS-203.
RX MEDLINE=94290509; PubMed=8019569;
RA Tuchman M., Piance R.J., Giguere Y., Lemieux B.;
RT "The ornithine transcarbamylase gene: new 'private' mutations in four
RN patients and study of a polymorphism.";
RN Hum. Mutat. 3:318-320(1994).
RN [19]
RP VARIANTS OTCD GLY-126; HIS-129 AND MET-172.
RX MEDLINE=94362715; PubMed=8081398;
RA Matsuura T., Hoshida R., Kiwaki K., Komaki S., Koike E., Endo F.,
RA Oyanagi K., Suzuki Y., Kato I., Ishikawa K., Yoda H., Kamitani S.,
RA Sakaki Y., Matsuda I.;
RT "Four newly identified ornithine transcarbamylase (OTC) mutations
RN (D126G, R129H, I172M and W322X) in Japanese male patients with early-
RN onset OTC deficiency.";
RN Hum. Mutat. 3:402-406(1994).
RN [20]
RP VARIANTS OTCD HIS-40; HIS-129; ARG-195; THR-225; GLN-277 AND
RN GLU-309 DEL.
RX MEDLINE=95038770; PubMed=7951259;
RA Tuchman M., Plante R.J., McCann M.T., Qureshi A.A.;
RT "Seven new mutations in the human ornithine transcarbamylase gene.";
RN Hum. Mutat. 4:57-60(1994).
RN [21]
RP VARIANTS OTCD THR-159 AND VAL-209.
RX MEDLINE=96070988; PubMed=8530002;
RA Garcia-Perez M.A., Sanjurjo P., Briones P., Garcia-Munoz M.J.,
RA Rubio V.;
RT "A splicing mutation, a nonsense mutation (Y167X) and two missense
RN mutations (I159T and A209V) in Spanish patients with ornithine
RN transcarbamylase deficiency.";
RN Hum. Genet. 96:549-551(1995).
RN [22]
RP VARIANTS OTCD GLU-269.
RX MEDLINE=96086561; PubMed=7474905;
RA Zimmer K.P., Matsuura T., Colombo J.-P., Koch H.G., Ullrich K.,
RA Deufel T., Harms E., Matsuda I.;
RT "A novel point mutation at codon 269 of the ornithine
RN transcarbamylase (OTC) gene causing neonatal onset of OTC
RN deficiency.";
RN J. Inher. Metab. Dis. 18:356-357(1995).
RN [23]
RP VARIANTS OTCD MET-125; ARG-188; VAL-209 AND LEU-302.
RX MEDLINE=96400964; PubMed=8807340;
RA Gilbert-Dussardier B., Segues B., Rozet J.-M., Rabier D., Calvas P.,
RA de Lumley L., Bonafant J.-P., Munnich A.;
RT "Partial duplication (dup. TCAC (178)) and novel point mutations
RN (T125W, G188R, A209V, and H302L) of the ornithine transcarbamylase
RN gene in congenital hyperammonemia.";
RN Hum. Mutat. 8:74-76(1996).
RN [24]
RP VARIANTS OTCD HIS-40; ASN-88; TYR-202 AND ASN-263.
RA Guardamagna O., Gatti E., Parini R., Plante R.J., Tuchman M.;
RT "Genotype-phenotype correlations in ornithine transcarbamylase
RN deficiency.";
RN Enzyme Protein 49:191-191(1996).
RN [25]
RP VARIANTS OTCD ASN-88; CYS-176; ALA-220; TYR-302 AND LYS-343.
RX MEDLINE=97114289; PubMed=8956038;

Query Match 65.0%; Score 141; DB 1; Length 354;
Best Local Similarity 59.0%; Pred. No. 1.7e-12;
Matches 29; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 MSLNRLILNKAALRAKHAFTSWVRNFRYKPVQSOVOLKPRDL 42
| | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 MLFNRLILNKAALRAKHAFTSWVRNFRYKPVQSOVOLKPRDL 42
| | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
IF2P_METAC
ID IF2P_METMA STANDARD; PRT; 591 AA.

Query Match 27.2%; Score 59; DB 1; Length 597;
Best Local Similarity 33.3%; Pred. No. 1.4;
Matches 14; Conservative 13; Mismatches 11; Indels 4; Gaps 2;

QY 3 SNLRILLNKAALRAKHAFTSWVRNFRYKPVQSOVOLKPRDL 42
| | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 255 ATLDVILYDGLTKGDTVIGSL--GEPIQIKVRAALKAPREL 294
| | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
IF2P_METMA
ID IF2P_METMA STANDARD; PRT; 591 AA.

```



```

ID AC RR2_LOTJA STANDARD; PRT; 236 AA.
DT 09BBS6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S2.
GN RPS2.
OS Lotus japonicus.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP "Complete structure of the chloroplast genome of a legume, Lotus
  japonicus."
RC STRAIN=Accession MG-20;
RC MEDLINE=21082929; PubMed=11214967;
RA Kato T., Kaneko T., Sato S., Nakamura Y., Tabata S.;
RT "Complete structure of the chloroplast genome of a legume, Lotus
  japonicus."
RL DNA Res. 7:323-330(2000).
CC -1- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AP002983; BAB33197.1;
CC HAMAP: MF_00291; 1.
CC InterPro: IPR001865; Ribosomal_S2.
CC Pfam: PF00318; Ribosomal_S2; 1.
CC PRINTS: PR00395; Ribosomal_S2.
CC TIGRFAMs: TIGR01011; rpsB_bact; 1.
CC PROSITE: PS00962; RIBOSOMAL_S2_1; 1.
CC PROSITE: PS00963; RIBOSOMAL_S2_2; 1.
CC Ribosomal protein; Chloroplast.
CC KW RIBOSOMAL_S2;
CC SEQUENCE 236 AA; 26982 MW; CE9238572325586 CRC64;
CC -----
Query Match 24.9%; Score 54; DB 1; Length 236;
Best Local Similarity 38.5%; Pred. No. 2.6;
Matches 10; Conservative 10; Mismatches 4; Indels 2; Gaps 1;

QY 2 LSNRLILNKA--ALRKATSMVRNF 25
Db 211 IASRLILNKLVAICEGSHSVIRNF 236
:::||||| |::|:|:|
11 IASRLILNKLVAICEGSHSVIRNF 236

RESULT 8
HEMN_SALTY STANDARD; PRT; 457 AA.
ID HEMN_SALTY
AC P37129;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Oxygen-independent coproporphyrinogen III oxidase (EC 1.1.1.11)
DE (Coproporphyrinogenase) (Coprogen oxidase).
GN HEMN OR STM4004 OR STY3877 OR r3617.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2;
RC MEDLINE=94252986; PubMed=8195073;
RA Xu K., Elliott T.;
RT "Cloning, DNA sequence, and complementation analysis of the
  Salmonella typhimurium hemN gene encoding a putative
  oxygen-independent coproporphyrinogen III oxidase."
  J. Bacteriol. 176:3196-3203(1994).
  [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RC MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
  LT2."
  Nature 413:852-856(2001).
RL Nature 413:852-856(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RC MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagers K.,
RA Krohn A., Larsen R.S., Leather S., Moule S., O'Garra P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
  enterica serovar Typhi CT18."
  Nature 413:848-852(2001).
RL Nature 413:848-852(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
RC MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
  and CT18."
  J. Bacteriol. 185:2330-2337(2003).
CC -1- FUNCTION: ANAEROBIC TRANSFORMATION OF COPROPORPHYRINOGEN-III INTO
  PROTOPORPHYRINOGEN-IX.
CC -1- COFACTOR: REQUIRES MAGNESIUM, ATP AND NAD (OR NADP) FOR ACTIVITY.
CC -1- PATHWAY: Porphyrin biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE ANAEROBIC COPROPORPHYRINOGEN III
  OXIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U06779; AAA19690.1;
CC EMBL: AE008887; AAL22843.1;
CC EMBL: AL627280; CAD03096.1;
CC EMBL: AE016846; AAO71118.1;
CC StyGene: SG10154; hemN
CC InterPro: IPR006638; Eip3.
CC InterPro: IPR004558; HemN.
CC Pfam: PF04055; Radical_SAM; 1.
CC SMART: SM00729; Eip3; 1.
CC TIGRFAMs: TIGR00538; hemN; 1.
CC Porphyrin biosynthesis; Oxidoreductase; Magnesium; NAD;
CC Complete proteome.
CC KW Complete;
CC SEQUENCE 457 AA; 52828 MW; 5667B4FE76204DAB CRC64;
CC -----
Query Match 24.9%; Score 54; DB 1; Length 457;
Best Local Similarity 26.5%; Pred. No. 5.3;
Matches 9; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

```



```

CGS5_YEAST          STANDARD;          PRT;      435 AA.
AC P30283;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE S-phase entry cyclin 5.
DE CLB5 OR YPR120C OR P9642.8.
GN Saccharomyces cerevisiae (Baker's yeast).
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BF264-15D;
RX MEDLINE=92387544; PubMed=1387626;
RA Epstein C.B., Cross F.R.
RT "CLB5: a novel B cyclin from budding yeast with a role in S phase.";
RL Genes Dev. 5:1695-1706(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB320;
RX MEDLINE=93307652; PubMed=8319908;
RA Schwob E., Nasmyth K.;
RT "CLB5 and CLB6, a new pair of B cyclins involved in DNA replication
in Saccharomyces cerevisiae.";
RL Genes Dev. 7:1160-1175(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=94074543; PubMed=8253070;
RA Kuehne C., Linder P.;
RT "A new pair of B-type cyclins from Saccharomyces cerevisiae that
function early in the cell cycle.";
RL EMBO J. 12:3437-3447(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313271; PubMed=9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansong W.,
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Dietrich F.S., Delius H., Dipaolo T., Dubois E., Duesterhoeft A.,
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
RA Hunnicke-Smith S., Hyman R., Johnston M., Kaiman S., Kleine K.,
RA Kemp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
RA Urestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA Zhong W.W., Zollner A., Vo D.H., Hani J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
RL Nature 387:103-105(1997).
CC -1- FUNCTION: REQUIRED FOR EFFICIENT PROGRESSION THROUGH S PHASE AND
CC WITH CDC28
CC -1- POSSIBLY FOR THE NORMAL PROGRESSION THROUGH MEIOSIS. INTERACTS
CC WITH CDC28
CC -1- DEVELOPMENTAL STAGE: MAXIMALLY EXPRESSED JUST BEFORE CELL CYCLE
CC START.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN AB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M91209; AAA34503.1; -
CC EMBL; X70435; CAA49893.1; -
CC
DR DR
DR EMBL; U40828; AAB68061.1; -
DR PIR; S31290; S31290.
DR SGD; S0006324; CLB5.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0000082; P:G1/S transition of mitotic cell cycle; IEP.
DR GO; GO:0000086; P:G2/M transition of mitotic cell cycle; IEP.
DR GO; GO:0006279; P:premeiotic DNA synthesis; IGI.
DR InterPro; IPR006670; Cyclin_Cterm.
DR InterPro; IPR004367; Cyclin_Nterm.
DR InterPro; IPR006671; Cyclin_N.
DR Pfam; PF00134; cyclin; 1.
DR Pfam; PF02984; cyclin_C; 1.
DR SMART; SM00385; CYCLIN; 2.
DR PROSITE; PS00292; CYCLINS; 1.
KW Cyclin. Cell cycle; Cell division; Multigene family.
SQ SEQUENCE 435 AA; 50431 MW; 5AD67EB841BA5759 CRC64;
Query Match 24.4%; Score 53; DB 1; Length 435;
Best Local Similarity 46.4%; Pred. No. 7;
Matches 13; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
QY 3 SNLRLLNKAAALRKAKTSMVNFYGGK 30
DB 28 SNLKILQNKRLSKNDSSKQKQVQDSKP 55
RESULT 12
EPF1_CHLMU
ID EPF1_CHLMU STANDARD; PRT; 185 AA.
AC Q9PKR6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation factor P 1 (EF-P 1).
GN EF1 OR TC0398.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Nigg;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Bass S.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Dodson R.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -1- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATES EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -1- PATHWAY: Protein biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the elongation factor P family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE002306; AAF39255.1; -
CC PIR; B81708; B81708.
CC TIGR; TC0398; -
CC HAMAP; MF_00141; -; 1.
CC InterPro; IPR001059; EF-P.

```





PIR; F85702; P85702.  
 PIR; G90844; G90844.  
 PIR; S25137; RGECCX.  
 ECGene; EG10646; narX.  
 InterPro; IPR003594; ATPbind\_ATPase.  
 InterPro; IPR003660; HAMP.  
 InterPro; IPR003661; His\_kinA.  
 InterPro; IPR005457; His\_kinase.  
 Pfam; PF00872; HATPase\_1.  
 Pfam; PF02518; HATPase\_C; 1.  
 SMART; SM00304; HAMP; 1.  
 SMART; SM00387; HATPase\_C; 1.  
 SMART; SM00388; Hiska; 1.  
 PROSITE; PS00885; HAMP; 1.  
 PROSITE; PS50109; HIS\_KIN; 1.  
 Sensory transduction; Transferrase; Kinase; Phosphorylation;  
 Transmembrane; Inner membrane; Nitrate assimilation;  
 Complete proteome.  
 FT DOMAIN 1 14  
 FT TRANSMEM 15 37  
 FT DOMAIN 38 151  
 FT TRANSMEM 152 174  
 FT DOMAIN 175 578  
 FT TRANSMEM 176 228  
 FT DOMAIN 393 587  
 FT MOD\_RES 399 399  
 FT CONFLICT 42 52  
 FT CONFLICT 374 374 A -> G (IN REF. 7).  
 SEQUENCE 598 AA; 67093 MW; 886BA0FC2C8F3CE CRC64;

Query Match 24.0%; Score 52; DB 1; Length 598;  
 Best Local Similarity 31.0%; Pred. NO. 14;  
 Matches 13; Conservative 9; Mismatches 18; Indels 2; Gaps 1;

QY 1 MLSNRLINKAAALKAHTSMVRNF--RYGKPVQSQVQLKRP 40  
 :|::||::||::||::||::||::||::||::||::||::||:  
 Db 448 LITFLQLTEPLGLPALEASCEYSARFGFPVKLDYLQLP 489

RESULT 14  
 POPB\_SCHPO STANDARD; PRT; 506 AA.  
 ID POFB\_SCHPO STANDARD; PRT; 506 AA.  
 AC Q09855; Q9P7V1.  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein pof11.  
 GN POF11 OR SPAC29E6.01 OR SPAC30.05.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SCHPO FROM N.A.  
 RC STRAIN=972.  
 RX MEDLINE=21668955; PubMed=11809834;  
 RA Katayama S., Kitamura K., Lehmann A., Nikaido O., Toda T.;  
 RT "Fission yeast F-box protein Pof3 is required for genome integrity and  
 telomere function.";  
 RL Mol. Biol. Cell 13:211-224(2002).  
 [2]  
 RP SCHPO FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., William R., Rajandream M.A., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,  
 RA Gentles S., Goblet A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Buckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

PIR; F85702; P85702.  
 DR PIR; G90844; G90844.  
 DR PIR; S25137; RGEENX.  
 DR Ecogene; EG10646; narX.  
 DR InterPro; IPR003594; ATPbind\_ATPase.  
 DR InterPro; IPR003660; HAMP.  
 DR InterPro; IPR003661; His\_kinA.  
 DR InterPro; IPR005457; His\_kinase.  
 DR Pfam; PF00872; HATPase\_1.  
 DR Pfam; PF02518; HATPase\_C; 1.  
 DR SMART; SM00304; HAMP; 1.  
 DR SMART; SM00387; HATPase\_C; 1.  
 DR SMART; SM00388; Hiska; 1.  
 DR PROSITE; PS00885; HAMP; 1.  
 DR PROSITE; PS50109; HIS\_KIN; 1.  
 KW Sensory transduction; Transferrase; Kinase; Phosphorylation;  
 KW Transmembrane; Inner membrane; Nitrate assimilation;  
 KW Complete proteome.  
 FT DOMAIN 1 14  
 FT TRANSMEM 15 37  
 FT DOMAIN 38 151  
 FT TRANSMEM 152 174  
 FT DOMAIN 175 578  
 FT TRANSMEM 176 228  
 FT DOMAIN 393 587  
 FT MOD\_RES 399 399  
 FT CONFLICT 42 52  
 FT CONFLICT 374 374 A -> G (IN REF. 7).  
 FT SEQUENCE 598 AA; 67093 MW; 886BA0FC2C8F3CE CRC64;

Query Match 24.0%; Score 52; DB 1; Length 598;  
 Best Local Similarity 31.0%; Pred. NO. 14;  
 Matches 13; Conservative 9; Mismatches 18; Indels 2; Gaps 1;

QY 1 MLSNRLINKAAALKAHTSMVRNF--RYGKPVQSQVQLKRP 40  
 :|::||::||::||::||::||::||::||::||::||::||:  
 Db 448 LITFLQLTEPLGLPALEASCEYSARFGFPVKLDYQLPPR 489

RESULT 14  
 POPB\_SCHPO STANDARD; PRT; 506 AA.  
 ID POBF\_SCHPO  
 AC Q09855; Q9P7V1;  
 DT 01-REB-1996 (Rel. 33, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein popfl.  
 GN POP11 OR SPAC29E6.01 OR SPAC30.05.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SCHPO  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=972;  
 RX MEDLINE=21668955; PubMed=11809834;  
 RA Katayama S., Kitamura K., Lehmann A., Nikaido O., Toda T.;  
 RT "Fission yeast F-box protein Popf3 is required for genome integrity and  
 telomere function.";  
 RL Mol. Biol. Cell 13:211-224(2002).  
 RN [2]  
 RP SCHPO  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Williams R., Rajandream M.A., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Buckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,



Search completed: July 24, 2003, 22:04:32  
Job time : 8.67857 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 22:02:24 ; Search time 14.3333 Seconds  
(without alignments)  
288.506 Million cell updates/sec

Title: US-08-765-244-22  
Perfect score: 217  
Sequence: 1 MLSNRLINKAALRAHIS.....NFRYKPVQSVQLKPRDLIC 43  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	199	91.7	354 1	OWRT
2	164	75.6	354 1	OWMS
3	141	65.0	354 1	OWHU
4	119	54.8	41 2	I52779
5	61	28.1	1286	B71413
6	60	27.6	351 2	F90409
7	57	26.3	354 2	JE0309
8	55.5	25.6	454 2	A86345
9	55	25.3	350 2	A48421
10	54	24.9	300 2	T32702
11	54	24.9	457 2	AD0950
12	54	24.9	1225 2	A56314
13	53.5	24.7	425 2	I40646
14	53	24.4	435 2	S31290
15	53	24.4	593 2	T47000
16	53	24.4	593 2	AB0239
17	52.5	24.2	185 2	B81708
18	52	24.0	213 2	B83182
19	52	24.0	335 2	T33457
20	52	24.0	598 1	RGEONX
21	52	24.0	598 2	F85702
22	52	24.0	598 2	G90844
23	51.5	23.7	260 2	G70233
24	51	23.5	506 2	T50211
25	51	23.5	598 2	AF0648
26	51	23.5	633 2	S76749
27	51	23.5	703 2	S84521
28	51	23.5	1779 2	T23130
29	50.5	23.3	138 2	F81900

short-chain-specif  
hypothetical prote  
hypothetical prote  
hypothetical prote  
motB homolog lafu  
CAAX prenyl protei  
probable molybdopt  
signal recognition  
calcium-dependent  
cytochrome c PA303  
purine NTPase [imp  
tumor suppressor p  
18c protein (clone  
hypothetical prote  
GTP-binding protei  
conserved hypothet

## ALIGNMENTS

### RESULT 1

OWRT  
ornithine carbamoyltransferase (EC 2.1.3.3) precursor - rat  
N:Alternate names: citrulline phosphorylase; ornithine transcarbamylase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 28-Feb-1986 #sequence revision 28-Feb-1986 #text change 11-Jun-1999  
C:Accession: A00563; A28042; A23090; S02466; I52976; I67609; I53457  
R:Rakiguchi, M.; Miura, S.; Mori, M.; Tatibana, M.; Nagata, S.; Kaziro, Y.  
Proc. Natl. Acad. Sci. U.S.A. 81, 7412-7416, 1984  
A:Title: Molecular cloning and nucleotide sequence of cDNA for rat ornithine carbamoyltransferase  
A:Reference number: A00563; MUID:85063800; PMID:6095294  
A:Accession: A00563  
A:Molecule type: mRNA  
A:Residues: 1-354 <TAK>  
A:Cross-references: GB:K03040; NID:9205873; PIDN:AAA41768.1; PID:9205874  
R:Rakiguchi, M.; Murakami, T.; Miura, S.; Mori, M.  
Proc. Natl. Acad. Sci. U.S.A. 84, 6136-6140, 1987  
A:Title: Structure of the rat ornithine carbamoyltransferase gene, a large, X chromo  
A:Reference number: A28042; MUID:87317609; PMID:3476935  
A:Accession: A28042  
A:Molecule type: DNA  
A:Residues: 1-354 <TAK>  
A:Cross-references: GB:M16933; GB:J02957; NID:9205884; PIDN:AAA41769.1; PID:9205886  
R:Kraus, J.P.; Hodges, P.E.; Williamson, C.L.; Horwich, A.L.; Kalousek, F.; Williams, Nucleic Acids Res. 13, 943-952, 1985  
A:Title: A cDNA clone for the precursor of rat mitochondrial ornithine transcarbamylase  
A:Reference number: A23090; MUID:85215524; PMID:3839075  
A:Accession: A23090  
A:Molecule type: mRNA  
A:Residues: 1-38, 'P', 40-240, 'S', 242-354 <KRA>  
A:Cross-references: GB:X01976  
R:Aoki, Y.; Sunaga, H.; Suzuki, K.T.  
Biochem. J. 250, 735-742, 1988  
A:Title: A cadmium-binding protein in rat liver identified as ornithine carbamoyltransferase  
A:Reference number: S02466; MUID:88268748; PMID:3390141  
A:Accession: S02466  
A:Molecule type: protein  
A:Residues: 33-56;293-302;307-317;322-329 <AOK>  
R:McIntyre, P.; Graf, L.; Mercer, J.F.B.; Wake, S.A.; Hudson, P.J.; Hoogenraad, N. DNA 4, 147-156, 1985  
A:Title: The primary structure of the imported mitochondrial protein, ornithine transcarbamoyltransferase  
A:Reference number: I52976; MUID:85203360; PMID:3838931  
A:Accession: I52976  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-354 <RES>  
A:Cross-references: GB:M1266; NID:9205871; PIDN:AAA41767.1; PID:9205872  
R:McIntyre, P.; Graf, L.; Mercer, J.; Peterson, G.; Hudson, P.J.; Hoogenraad, N. FEBS Lett. 177, 41-46, 1984  
A:Title: A highly basic N-terminal extension of the mitochondrial matrix enzyme ornithine carbamoyltransferase  
A:Reference number: I53457; MUID:85051832; PMID:6548714  
A:Accession: I67609

A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-102 <RE2>  
 A:Cross-references: EMBL:X01178; NID:g56802; PIDN:CAA25618.1; PID:g56803  
 A:Accession: I53457  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-43, 'N', 45-99, 'R', 101-102 <RE3>  
 A:Cross-references: GB:K03041; NID:g205889; PIDN:AAA41771.1; PID:g205890  
 C:Genetics:  
 A:Introns: 26/2; 72/3; 100/1; 129/2; 180/3; 221/3; 239/3; 289/3; 335/3  
 C:Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase  
 C:Keywords: arginine biosynthesis; homotrimer; mitochondrion; transferase; urea cycle  
 F:1-32/Domain: transit peptide (mitochondrion) #status predicted <TNP>  
 F:33-354/Product: ornithine carbamoyltransferase #status predicted <NAT>  
 F:40-342/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>

Query Match 91.78; Score 199; DB 1; Length 354;  
 Best Local Similarity 97.68; Pred. No. 9.3e-20;  
 Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSNRLILLNKAALRKHAHTSMVRNFRYKPKVQSQVQLKPRDL 42  
 DB 1 MLSNRLILLNKAALRKHAHTSMVRNFRYKPKVQSQVQLKGRDL 42

## RESULT 2

## OWHU

N:Alternate names: citrulline phosphorylase; ornithine transcarbamylase  
 C:Species: Mus musculus (house mouse)  
 C:Date: 31-Mar-1993 #sequence, revision 13-Mar-1997 #text\_change 11-Jun-1999  
 A:Accession: A43609; S03407; I55252  
 R:Veres, G.; Gibbs, R.A.; Scherer, S.E.; Caskey, C.T.  
 Science 237, 415-417, 1987  
 A:Title: The molecular basis of the sparse fur mouse mutation.  
 A:Reference number: A43609; MUID:87263407; PMID:3603027  
 A:Accession: A43609  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-354 <VER>  
 A:Cross-references: GB:ML7030; NID:g200162; PIDN:AAA39865.1; PID:g200163  
 R:Scherer, S.E.; Veres, G.; Caskey, C.T.  
 Nucleic Acids Res. 16, 1593-1601, 1988  
 A:Title: The genetic structure of mouse ornithine transcarbamylase.  
 A:Reference number: S03407; MUID:88157177; PMID:2831503  
 A:Accession: S03407  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-194, 'R', 196-335 <SCH>  
 A:Cross-references: EMBL:X07092  
 R:Veres, G.; Craigen, W.J.; Caskey, C.T.  
 J. Biol. Chem. 261, 7588-7591, 1986  
 A:Title: The 5' flanking region of the ornithine transcarbamylase gene contains DNA sequences that are essential for expression in transgenic mice.  
 A:Reference number: I55252; MUID:86224037; PMID:3011788  
 A:Accession: I55252  
 A:Status: translation not shown; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-19, 'LLWFDF', <RES>  
 A:Cross-references: GB:ML2176; NID:g200160; PIDN:AAA39864.1; PID:g554248  
 A:Note: The end of this sequence is near the boundary of the cloned region and may be an artifact.  
 C:Genetics:  
 A:Gene: OTC  
 A:Map position: X  
 A:Introns: 26/2; 72/3; 100/1; 129/2; 180/3; 221/3; 239/3; 289/3  
 C:Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase  
 C:Keywords: mitochondrion; transferase; urea cycle  
 F:1-32/Domain: transit peptide (mitochondrion) #status predicted <TNP>  
 F:33-335/Product: ornithine carbamoyltransferase #status predicted <NAT>  
 F:40-342/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>

Query Match 75.68; Score 164; DB 1; Length 354;  
 Best Local Similarity 83.38; Pred. No. 6.6e-15;

Matches 35; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MLSNRLILLNKAALRKHAHTSMVRNFRYKPKVQSQVQLKPRDL 42  
 DB 1 MLSNRLILLNKAALRKHAHTSMVRNFRYKPKVQSQVQLKGRDL 42  
 RESULT 3  
 OWHU  
 N:Alternate names: citrulline phosphorylase; ornithine transcarbamylase  
 C:Species: Homo sapiens (man)  
 C:Date: 28-Feb-1986 #sequence, revision 31-Mar-1993 #text\_change 16-Jun-2000  
 C:Accession: A41444; B41444; A00562; I38078; JC4672; I59039; I54377  
 R:Hata, A.; Tsuzuki, T.; Shimada, K.; Takiguchi, M.; Mori, M.; Matsuda, I.  
 J. Biochem. 103, 302-308, 1988  
 A:Title: Structure of the human ornithine transcarbamylase gene.  
 A:Reference number: A41444; MUID:88227905; PMID:2836378  
 A:Accession: A41444  
 A:Molecule type: DNA  
 A:Residues: 1-354 <HAT>  
 A:Cross-references: GB:D00230; NID:g219957; PIDN:BAA00161.1; PID:g219959  
 A:Accession: B41444  
 A:Molecule type: mRNA  
 A:Residues: 1-354 <HA2>  
 A:Cross-references: GB:D00230; NID:g219957; PIDN:BAA00161.1; PID:g219959  
 R:Horwich, A.L.; Fenton, W.A.; Williams, K.R.; Kalousek, F.; Kraus, J.P.; Doolittle, R.  
 Science 224, 1068-1074, 1984  
 A:Title: Structure and expression of a complementary DNA for the nuclear coded precursor of ornithine carbamoyltransferase.  
 A:Reference number: A00562; MUID:84196410; PMID:6372096  
 A:Accession: A00562  
 A:Molecule type: mRNA  
 A:Residues: 1-100, 'F', 102-110, 'P', 112-192, 'CF', 195-269, 'R', 271-354 <HOR>  
 A:Cross-references: GB:D00230  
 R:Hata, A.; Tsuzuki, T.; Shimada, K.; Takiguchi, M.; Mori, M.; Matsuda, I.  
 J. Biochem. 100, 717-725, 1986  
 A:Title: Isolation and characterization of the human ornithine transcarbamylase gene.  
 A:Reference number: I38078; MUID:87057134; PMID:3782067  
 A:Accession: I38078  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-26 <RES>  
 A:Cross-references: EMBL:X04443; NID:g35162; PIDN:CAA28039.1; PID:g35163  
 R:Wheeler, V.C.; Prodromou, C.P.; Pearl, L.H.; Williamson, R.; Coutelle, C.  
 Gene 169, 251-255, 1996  
 A:Title: Synthesis of a modified gene encoding human ornithine transcarbamylase for eukaryotic expression.  
 A:Reference number: JC4672; MUID:96194812; PMID:8647457  
 A:Accession: JC4672  
 A:Molecule type: DNA  
 A:Residues: 'M', 33-100, 'F', 102-110, 'P', 112-192, 'CF', 195-269, 'R', 271-354 <WHE>  
 A:Note: This report represents a synthetic gene designed for expression in (rather than in a prokaryotic host).  
 R:Horwich, A.L.; Kalousek, F.; Rosenberg, L.E.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 4930-4933, 1985  
 A:Title: Arginine in the leader peptide is required for both import and proteolytic cleavage of the human ornithine carbamoyltransferase.  
 A:Reference number: I59039; MUID:85270440; PMID:3895227  
 A:Accession: I59039  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-36 <RE2>  
 A:Cross-references: GB:ML1235; NID:g189408; PIDN:AAA59976.1; PID:g189409  
 R:Gilbert-Dussardier, B.; Rabier, D.; Strautnieks, S.; Segues, B.; Bonnefont, J.P.; M. Hum. Mol. Genet. 3, 831-832, 1994  
 A:Title: A novel arginine (245) to glutamine change in exon 8 of the ornithine carbamoyltransferase gene.  
 A:Reference number: I54377; MUID:94362689; PMID:8081373  
 A:Accession: I54377  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 269-276, 'Q', 278-289 <RE3>  
 A:Cross-references: GB:S73640; NID:g688001; PIDN:AB31859.1; PID:g688002  
 A:Note: This sequence represents a disease defect in ornithine carbamoyltransferase. The active enzyme is a dimer of identical chains with one tightly bound zinc ion per subunit.

[illegible]

ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: A86345

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-454 <STO>

A:Cross-references: GB:AE005172; NID:g8920642; PIDN:AAF81364.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 25.6%; Score 55.5; DB 2; Length 454;

Best Local Similarity 45.2%; Pred. No. 9.6;

Matches 14; Conservative 5; Mismatches 11; Indels 1; Gaps 1;

QY 11 KAALRAHT-SMVRNFRYKPKVQSQVQLKPR 40

DB 422 KAAYSTSDSPRNRESQPMGSPVOARPR 452

RESULT 9

A48421

ornithine transcarbamylase - bullfrog

C:Species: Rana catesbeiana (bullfrog)

C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Jun-1999

C:Accession: A48421

R:Helbing, C.; Gergely, G.; Atkinson, B.G.

Dev. Genet. 13, 289-301, 1992

A:Title: Sequential up-regulation of thyroid hormone beta receptor, ornithine transcarbamylase and thyroid hormone-induced metamorphosis.

A:Reference number: A48421; MUID:93177976; PMID:1291156

A:Accession: A48421

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-350 <HEL>

A:Cross-references: GB:M95193; NID:g213683; PIDN:AAA49528.1; PID:g213684

A:Experimental source: liver

A:Note: sequence extracted from NCBI backbone (NCBIN:126154, NCBI:P126155)

C:Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase

C:Keywords: mitochondrion

F:36-338/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>

Query Match 25.3%; Score 55; DB 2; Length 350;

Best Local Similarity 38.1%; Pred. No. 8.5;

Matches 16; Conservative 9; Mismatches 13; Indels 4; Gaps 2;

QY 1 MSLNRLILNKAALRAHTSMVRNFRYKPKVQSQVQLKPRDL 42

DB 1 MLHMTIIN-ASWRYGNRCIVRQFGS---QTSQLKGRDL 38

RESULT 10

T32702

hypothetical protein C14C6.13 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 04-Mar-2000

C:Accession: T32702

R:David, M.; Wohldmann, P.; Bauer, C.; Antoniou, B.

submitted to the EMBL Data Library, December 1997

A:Description: The sequence of C. elegans cosmid C14C6.

A:Reference number: 221210

A:Accession: T32702

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-300 <DAB>

A:Cross-references: EMBL:AF039051; PIDN:AAB94269.1; GSPDB:GN00023; CESP:C14C6.13

A:Experimental source: strain Bristol N2; clone C14C6

C:Genetics:

A:Gene: CESP:C14C6.13

A:Map position: 5

A:Introns: 91/1; 129/2; 145/3; 267/3

C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.13

Query Match 24.9%; Score 54; DB 2; Length 300;

Best Local Similarity 31.4%; Pred. No. 9.9;

Matches 11; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 2 LSNRLILNKAALRAHTSMVRNFRYKPKVQSQVQ 36

DB 199 LDNIRELCQSALQKSETRPNTGTFYRINSEIQ 233

RESULT 11

AD0950

oxygen-independent coproporphyrinogen III oxidase [imported] - Salmonella enterica subsp. enterica serovar Typhi

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C:Accession: AD0950

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church

th, T.; Conington, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Fair

, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se

A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AD0950

A:Molecule type: DNA

A:Residues: 1-457 <PAR>

C:Cross-references: GB:AL513382; PIDN:CAD03096.1; PID:g16504733; GSPDB:GN00176

C:Genetics:

A:Gene: STY3877

C:Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 24.9%; Score 54; DB 2; Length 457;

Best Local Similarity 26.5%; Pred. No. 16;

Matches 9; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 9 LNKALRAHTSMVRNFRYKPKVQSQVQLKPRDL 42

DB 118 LNKAOISRLMTLLRNHFNTDAISIEVDPREI 151

RESULT 12

A56514

chromokinesin - chicken

C:Species: Gallus gallus (chicken)

C:Date: 21-Jul-1995 #sequence\_revision 28-Jul-1995 #text\_change 02-Feb-2001

C:Accession: A56514; I50691

R:Wang, S.Z.; Adler, R.

J. Cell Biol. 128, 761-768, 1995

A:Title: Chromokinesin: a DNA-binding, kinesin-like nuclear protein.

A:Reference number: A56514; MUID:95181533; PMID:7876303

A:Accession: A56514

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1225 <WAN>

A:Cross-references: GB:U18309; NID:g603760; PIDN:AAC59666.1; PID:g603761

R:Wang, S.Z.; Adler, R.

Proc. Natl. Acad. Sci. U.S.A. 91, 1351-1355, 1994

A:Title: A developmentally regulated basic-leucine zipper-like gene and its expressio

A:Reference number: A53451; MUID:94151328; PMID:8108415

A:Accession: I50691

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 728-1086, 'RI' <WA2>

A:Cross-references: EMBL:U04821; NID:g440792; PIDN:AAA18960.1; PID:g440793

C:Genetics:

A:Gene: sw3-3



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 21:59:18 ; Search time 37.881 Seconds  
(without alignments)  
180.176 Million cell updates/sec

Title: US-08-765-244-22  
Perfect score: 217  
Sequence: 1 MSLNLRILLKALRAKHAHS.....NFRYCKPVQSVQLKPRDLIC 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	217	100.0	43	17	AA90584
2	159	73.3	32	23	ABG30857
3	109	50.2	258	21	AA18445
4	108	49.8	32	23	ABG30856
5	105	48.4	32	22	AA64224
6	92	42.4	32	15	AA48260
7	92	42.4	32	20	AAW6358
8	92	42.4	32	23	ABG2993
9	92	42.4	32	24	ABP56588

10	79.5	36.6	31	21	AA15704
11	79.5	36.6	31	21	AA22835
12	56.5	26.0	149	22	AAU23468
13	54	24.9	866	22	AB57741
14	53.5	24.7	755	22	AB59259
15	52	24.0	165	20	AA04933
16	52	24.0	774	22	ABG29304
17	52	24.0	1124	22	AB59241
18	51.5	23.7	138	21	AA75465
19	51	23.5	1169	17	AA96126
20	50.5	23.3	138	21	AA75466
21	50.5	23.3	296	22	AB64170
22	50.5	23.3	2379	23	ABP62760
23	50	23.0	564	23	ABP73265
24	50	23.0	746	22	AA93938
25	50	23.0	975	22	AA94042
26	50	23.0	975	23	AA018173
27	49	22.6	74	22	AAU60044
28	49	22.6	457	16	AA78184
29	49	22.6	624	23	ABR98858
30	48.5	22.4	75	22	AA64583
31	48.5	22.4	75	22	AA64583
32	48.5	22.4	75	22	AA64583
33	48.5	22.4	185	20	AA37084
34	48.5	22.4	285	22	AA33373
35	48.5	22.4	334	22	AA33205
36	48.5	22.4	684	22	AA989868
37	48.5	22.4	4987	12	AA07559
38	48.5	22.4	4987	12	AA07559
39	48.5	22.4	5035	13	AA25450
40	48.5	22.4	5072	12	AA11510
41	48.5	22.4	5081	22	ABG04969
42	48.5	22.4	5081	22	AB11480
43	48	22.1	13	21	AA57007
44	48	22.1	93	22	AB10986
45	48	22.1	96	21	AA603218

## ALIGNMENTS

### RESULT 1

AA90584

ID AA90584 standard; Protein; 43 AA.

XX AA90584;

XX 25-MAR-2003 (updated)

DT 31-OCT-1996 (first entry)

XX DE

DE Rat ornithine transcarbamylase signal peptide.

XX promoter; peptide-nucleic acid; cyclised; gene therapy; target;

XX site-directed mutagenesis; introductin; protein transport.

XX OS Synthetic.

XX DE19520815-A1.

XX PD 21-DEC-1995.

XX PF 11-JUN-1995; 95DE-1020815.

XX PR 16-JUN-1994; 94DE-4421079.

XX (SEIB/) SEIBEL P.

XX Seibel A, Seibel P;

XX WPI; 1996-041226/05.

XX Replicable and transcriptionally active plasmid carrying signal

PT peptide for specific target - useful for site directed mutagenesis

1	217	100.0	43	17	AA90584
2	159	73.3	32	23	ABG30857
3	109	50.2	258	21	AA18445
4	108	49.8	32	23	ABG30856
5	105	48.4	32	22	AA64224
6	92	42.4	32	15	AA48260
7	92	42.4	32	20	AAW6358
8	92	42.4	32	23	ABG2993
9	92	42.4	32	24	ABP56588

PT and molecular therapy of genetic diseases.  
XX Disclosure; Column 11; 24pp; German.  
XX Two modified oligonucleotides (introducing PstI and XhoI sites) were  
CC used to amplify a region of the human mitochondrial (mt) genome contg.  
CC the light strand promoter, mtDNA ori of the heavy strand, CSP/s  
CC ("conserved sequence blocks"), and a regulation site for DNA replication.  
CC Behind this fragment (5' direction) a synthetic multiple cloning site  
CC was introduced, generating a product with overhang ends. The synthetic  
CC region also introduced a bidirectional mt transcription termination  
CC sequence. The amplification product, synthetic fragment and pBluescript  
CC were ligated and recombinant plasmid 1 (AAR12315) was produced. Human mt  
CC 16S rRNA (differing from the native RNA only in having a modified  
CC nucleotide) was isolated by PCR from chloramphenicol resistant HeLa  
CC cells and inserted into plasmid 1 to form plasmid 2 (AAR12316). The  
CC cloned insert was isolated as a BsaI fragment and cyclised using hairpin  
CC loop oligonucleotides, one of which carried the required signal peptide  
CC (the present sequence). The cyclised product was purified by treatment  
CC with exonuclease III. In a modification, the signal peptide was attached  
CC after cyclisation. The new plasmids were able to impart chloramphenicol  
CC resistance to otherwise sensitive B lymphocytes and fibroblasts.  
CC Similar plasmids without a signal peptide could not do this. The  
CC plasmids can be introduced into eukaryotic cells, esp. for site-directed  
CC mutagenesis or molecular therapy of genetic diseases, targeting nucleic  
CC acid in cells or their organelles via the protein transport route.  
XX (Updated on 25-MAR-2003 to correct PR field.)  
XX  
SQ Sequence 43 AA;  
Query Match 100.0%; Score 217; DB 17; Length 43;  
Best Local Similarity 100.0%; Pred. No. 3.1e-26;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLSNRLILLNKAALRKAHTSMVRNFRYKPGVQSLKPRDLC 43  
|||||  
DB 1 MLSNRLILLNKAALRKAHTSMVRNFRYKPGVQSLKPRDLC 43  
|||||  
RESULT 2  
ABG30857  
ID ABG30857 standard; Peptide; 32 AA.  
XX  
AC ABG30857;  
XX  
DT 07-OCT-2002 (first entry)  
XX  
DE Rat ornithine transcarbamylase signal peptide.  
XX  
KW Rat; signal peptide; ornithine transcarbamylase; MOT;  
KW recombinant vector; fusion protein; extranuclear gene.  
XX  
OS Rattus sp.  
XX  
PN JP2002176988-A.  
XX  
PD 25-JUN-2002.  
XX  
PF 14-DEC-2000; 2000JP-0380975.  
XX  
PR 14-DEC-2000; 2000JP-0380975.  
XX  
PA (TANA/) TANAKA M.  
PA (GIFU-) ZH GIFUKEN KOKUSAI BIO KENYUJO.  
PA (OYOS-) ZH OYO SEIKAGAKU KENKYUSHO.  
XX  
DR WPI; 2002-569946/61.  
DR N-PSDB; ABK68419.  
XX  
XX A recombinant vector for expressing a fused protein, useful for  
PT decomposing an extranuclear gene of a nonhuman organism -  
XX  
PS Disclosure; Page 3; 15pp; Japanese.

XX The invention relates to a recombinant vector for expressing a fused  
CC protein containing a fused gene in which a base sequence defining a  
CC transfer signal peptide to small cellular organs having an extranuclear  
CC gene is combined with a base sequence defining the amino acid sequence of  
CC a restriction enzyme recognising a defined base sequence. The vector is  
CC used for decomposing an extranuclear gene of a nonhuman organism.  
CC The present sequence represents the signal peptide of rat mitochondrial  
CC ornithine transcarbamylase (MOT) which may be used in the vector of the  
XX invention.  
SQ Sequence 32 AA;  
Query Match 73.3%; Score 159; DB 23; Length 32;  
Best Local Similarity 100.0%; Pred. No. 2.4e-17;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLSNRLILLNKAALRKAHTSMVRNFRYKPGVQ 32  
|||||  
DB 1 MLSNRLILLNKAALRKAHTSMVRNFRYKPGVQ 32  
|||||  
RESULT 3  
AAB18445  
ID AAB18445 standard; Protein; 258 AA.  
XX  
AC AAB18445;  
XX  
DT 15-JAN-2001 (first entry)  
XX  
DE Protein encoded by plasmid pUOATP2.  
XX  
KW Mitochondrial genome; mitochondrial encephalomyopathy; lactic acidosis;  
KW Leber hereditary optic neuropathy; myoclonic epilepsy;  
KW neurogenic muscular weakness; ataxia; retinitis pigmentosa;  
KW Kearns-Sayre syndrome; Leigh syndrome; Pearson Marrow pancreas syndrome;  
KW aminoglycoside-associated deafness; diabetes; deafness; leukodystrophy;  
KW hypotonia; autism; sudden infant death syndrome; hypoglycemia; leukaemia;  
KW thrombocytopenia; migraine; hearing loss; stroke;  
KW refractory infantile reflux; carnitine deficiency; multiple sclerosis;  
KW blindness; optic atrophy; renal tubular acidosis; cardiomyopathy;  
KW chronic pancreatitis; ATPase 6.  
XX  
OS Synthetic.  
XX  
PN WO200053773-A2.  
XX  
PD 14-SEP-2000.  
XX  
PF 08-MAR-2000; 2000WO-US06037.  
XX  
PR 08-MAR-1999; 99US-0123336.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Zullo SJ, Eisenstadt JM;  
XX  
DR WPI; 2000-565602/52.  
DR N-PSDB; AAA75084.  
XX  
XX Functionally complementing one or more defects, mutations, or deletions  
PT in a mitochondrial genome of a cell, useful for treating diabetes with  
PT deafness, comprises introducing mitochondrial DNA into the nuclear  
PT genome -  
XX  
PS Example 1; Fig 3; 49pp; English.  
XX  
XX The specification describes a method for functionally complementing  
CC one or more defects, mutations, or deletions in a mitochondrial genome of  
CC a cell having a nuclear genome. The method is used for treating a disease  
CC or disorder that arises from deletion of the protein-encoding genes of  
CC the mitochondrial genome. Alternatively, the method is used for treating  
CC a disease or disorder that arises from one or more defects, deletions or



DT 29-JUL-1994 (first entry)

1 MLENRLILLNNAAPRNGHNEMVRNFRCGOPL 32

DT 29-JUL-1994 (first entry)

XX DE Mitochondrial matrix retention signal.

XX KW Single chain antibody; sFv; heavy chain; light chain;

XX KW variable domain; hydrophilic linker; antibodies; targeting;

XX KW subcellular localisation signal; mitochondrial matrix;

XX KW retention signal.

XX OS Synthetic.

XX FT Key Location/Qualifiers

XX FT Misc-difference 7 /note= "not defined"

XX FT Misc-difference 8 /note= "not defined"

XX FT Misc-difference 32 /note= "not defined"

XX FT WO9402610-A1.

XX PN 03-FEB-1994.

XX PD 16-JUL-1993; 93WO-US06735.

XX PF 17-JUL-1992; 92US-0916939.

XX PR 17-MAR-1993; 93US-0045274.

XX XX (DAND ) DANA FARBER CANCER INST INC.

XX PA Haseltine WA, Marasco WA;

XX PI WPI; 1994-048868/06.

XX DR Intracellular binding of antigens - by using antibody targeting

XX PT with vector system, for e.g. tumour suppression

XX PS Disclosure; Page 103; 155pp; English.

XX CC New vector systems comprise a sequence adapted for intracellular

XX CC delivery and expression contg. a promoter operably linked to an

XX CC antibody gene encoding an antibody which binds to a specific target

XX CC antigen. The antibody is esp. a single chain antibody in which the

XX CC heavy and light chain variable regions are joined via a hydrophilic

XX CC linker peptide. Localisation sequences are pref. included in the

XX CC constructs. The sequence AAR48260 is a mitochondrial matrix retention

XX CC signal.

XX CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 32 AA;

Query Match 42.4%; Score 92; DB 15; Length 32;

Best Local Similarity 61.3%; Pred. No. 6.8e-07;

Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAAALRKAHTSMVRNFRYGRPV 31

Db 1 MLFNLRXXLLNNAAFRGHNMVNRFCGQPL 31

RESULT 7

AAW96358

ID AAW96358 standard; Peptide; 32 AA.

XX AC AAW96358;

XX AC

XX DT 19-JUL-1999 (first entry)

XX DE Mitochondrial matrix localisation signal peptide.

XX KW Antibody; immune response; modulation; MHC; IRM; receptor;

XX KW intrabody; major histocompatibility complex; graft rejection;

XX KW immunomodulatory response molecule; regulation; transplantation;

XX KW retention signal; localisation signal; golgi apparatus; ER;

KW endoplasmic reticulum.

XX Synthetic.

XX OS

XX FT Key Location/Qualifiers

XX FT Misc-difference 7 /note= "Any amino acid"

XX FT Misc-difference 8 /note= "Any amino acid"

XX FT Misc-difference 32 /note= "Any amino acid"

XX FT WO9914353-A2.

XX PN 25-MAR-1999.

XX PD 18-SEP-1998; 98WO-US19563.

XX PF 19-SEP-1997; 97US-0059339.

XX PR (DAND ) DANA FARBER CANCER INST INC.

XX PA Marasco W, Mhashikar A;

XX PI WPI; 1999-229546/19.

XX DR Altering the regulation of the immune system

XX PT Disclosure; Page 28; 56pp; English.

XX PS

XX CC Intracellular binding to a desired target by an intracellularly

XX CC expressed antibody (i.e. an intrabody) can be used to knock out

XX CC multiple locuses of immunomodulatory receptor molecules (IRMs),

XX CC so that the expression of multiple major histocompatibility (MHC)

XX CC molecules is blocked. This selective targeting of IRMs, their

XX CC pathways or components, can be used to selectively regulate the

XX CC immune system by controlling expression of these molecules and

XX CC preventing an undesired immune response in a cell. Any component

XX CC of the MHC pathway or the MHC assembly line or antigen presentation

XX CC can be targeted. Intrabodies can be used to knock out the immune

XX CC response in a particular tissue or portion of the body to prepare

XX CC it for cell or tissue transplantation. Alternatively, an organ for

XX CC transplantation can be perfused with the intrabody ex vivo. The

XX CC intrabodies can comprise whole antibodies, heavy chains, Fab'

XX CC fragments, single-chain antibodies and diabodies. The intrabodies

XX CC also comprise an intracellular localisation signal to facilitate

XX CC interception of expressed proteins. For example, if the target was

XX CC a cell surface receptor, the antibody would comprise a leader

XX CC sequence and an endoplasmic reticulum (ER) or Golgi apparatus

XX CC retention signal. This peptide is a localisation sequence for the

XX CC mitochondrial matrix. For other localisation sequences see

XX CC AAW96345-W96377.

XX SQ Sequence 32 AA;

Query Match 42.4%; Score 92; DB 20; Length 32;

Best Local Similarity 61.3%; Pred. No. 6.8e-07;

Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAAALRKAHTSMVRNFRYGRPV 31

Db 1 MLFNLRXXLLNNAAFRGHNMVNRFCGQPL 31

RESULT 8

ABG92993

ID ABG92993 standard; Peptide; 32 AA.

XX AC ABG92993;

XX AC

XX DT 20-NOV-2002 (first entry)

XX DE Localisation sequence to direct antibodies to the mitochondria.

XX Regulator; transcription; cell death; phenotype; molecular scaffold;  
KW gene therapy; cancer; cardiovascular disease; arrhythmia; heart failure;  
KW ischaemia; obesity; neurodegenerative disease; Alzheimer's disease;  
KW bone pathology; dermatologic disease; psoriasis; infection; AIDS;  
KW acquired immunodeficiency syndrome; cosmetic; wound healing;  
KW antibiotic transport; drug toxicity; drug resistance; immunobiology;  
KW inflammation; allergic response; human immunodeficiency virus.  
XX Unidentified.  
XX WO200262822-A2.  
XX PD 15-AUG-2002.  
XX PF 04-FEB-2002; 2002WO-US02814.  
XX PR 02-FEB-2001; 2001US-265589P.  
XX PR 05-FEB-2001; 2001US-265880P.  
XX PR 27-FEB-2001; 2001US-271423P.  
XX PR 23-JAN-2001; 2001US-263226P.  
XX PR 28-MAR-2000; 2000US-192586P.  
XX PR 22-SEP-1997; 97US-935377P.  
XX PA (UYRP ) UNIV ROCHESTER.  
XX PI Zauderer M, Smith ES;  
XX WPI; 2002-643398/69.  
XX PT Identifying regulator polypeptides which influence target  
XX PT transcriptional regulatory regions, useful for treating cancer,  
XX PT comprises introducing host cells expressing the polypeptide into a  
XX PT library of polynucleotides  
XX PS Disclosure; Page 37; 224pp; English.  
XX The invention discloses a method for identifying polynucleotides encoding  
XX a regulator polypeptide, whose expression induces activation of a target  
XX transcriptional regulatory region in a host cell. The method comprises  
XX providing a population of eukaryotic host cells capable of expressing the  
XX polypeptide, introducing into the host cell a library of polynucleotides  
XX encoding the polypeptides, permitting expression of the polypeptides and  
XX then recovering them from the host cells. The target transcriptional  
XX regulatory region is operably associated with a polynucleotide encoding a  
XX gene product, the expression of which results in host cell death or cause  
XX the host cells to exhibit a pre-determined modified phenotype and where  
XX the gene product is expressed upon activation of target transcriptional  
XX regulatory region. Each candidate regulator polypeptide comprises a  
XX candidate peptide and a molecular scaffold fused to the peptide so that  
XX the peptide is displayed on the surface of the candidate regulator  
XX polypeptide. The methods are useful in selecting and/or screening  
XX regulator molecules, such as polypeptides, which directly or indirectly  
XX induce or suppress the transcriptional activation of a target  
XX transcriptional regulatory region in a eukaryotic host cell. These  
XX regulator molecules may be used (e.g. in gene therapy) for preventing or  
XX treating cancers (e.g. breast or ovarian cancer), cardiovascular diseases  
XX (e.g. arrhythmia, heart failure, ischaemia), obesity, neurodegenerative  
XX diseases (e.g. Alzheimer's disease), bone pathologies, dermatologic  
XX diseases (e.g. psoriasis), infections (e.g. viral, bacterial), acquired  
XX immunodeficiency syndrome (AIDS). In cosmetic applications and in wound  
XX healing. The method is also useful in screening regulator molecules that  
XX block antibiotic transport mechanisms, in drug toxicities and drug  
XX resistance applications and in improving the performance of existing or  
XX developmental drugs. It may also be used in immunobiology, inflammation,  
XX allergic response and in biotechnology applications. The sequences  
XX presented in ABG92946-ABG93029 are examples of regulator polypeptides.  
XX Sequence 32 AA;  
XX Query Match 42.4%; Score 92; DB 23; Length 32;  
XX Best Local Similarity 61.3%; Pred. No. 6.8e-07;  
XX Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MSLNRLILLNKAALRKAFTSMVRNFRYKPV 31  
II III II II II IIII I:  
DB 1 MLENLRLXLLNNAAPRHGHENFVRNFRGQPL 31  
RESULT 9  
ABP56588  
ID ABP56588 standard; Peptide; 32 AA.  
XX AC ABP56588;  
XX 24-MAR-2003 (first entry)  
XX DE Mitochondrial matrix targeting peptide SEQ ID NO:54.  
XX KW Identification; intrabody; eukaryotic cell; immunoglobulin; selection;  
KW cardiovascular; diminished arrhythmia potential; cardiomyocyte; stroke;  
KW enhanced contractile property; heart failure; arrhythmia; embolic;  
KW sarcolemmal calcium cycling; artery; arteriole; angina; atherosclerosis;  
KW LDL metabolism; HDL metabolism; skin biology; keloid formation.  
XX OS Unidentified.  
XX Key Location/Qualifiers  
FH Misc-difference 7 /note= "any amino acid"  
FT Misc-difference 8 /note= "any amino acid"  
FT Misc-difference 32 /note= "any amino acid"  
FT Misc-difference 32 /note= "any amino acid"  
XX WO200286096-A2.  
XX PD 31-OCT-2002.  
XX PF 23-JAN-2002; 2002WO-US01677.  
XX PR 23-JAN-2001; 2001US-263225P.  
XX PR 24-JAN-2001; 2001US-263220P.  
XX PR 27-FEB-2001; 2001US-271422P.  
XX PR 15-JUN-2001; 2001US-298095P.  
XX (UYRP ) UNIV ROCHESTER MEDICAL CENT.  
XX Zauderer M, Wei C, Smith ES;  
XX WPI; 2003-103408/09.  
XX Selecting polynucleotides encoding an intracellular immunoglobulin  
XX which induces a modified phenotype in a eukaryotic host cell, by  
XX introducing library of polynucleotides encoding immunoglobulin subunit  
XX polypeptides -  
XX Disclosure; Page 44; 257pp; English.  
XX The present invention describes a method for selecting polynucleotides  
XX (PNS) encoding an intracellular immunoglobulin molecule or its fragment  
XX whose expression induces a modified phenotype in a eukaryotic host cell  
XX (I). The method comprises introducing into (I) a first and second library  
XX of PNS encoding, through operable association with a transcriptional  
XX control region, first and second intracellular immunoglobulin subunit  
XX polypeptides, respectively. The method is useful for selecting  
XX polynucleotides which encode an intracellular immunoglobulin molecule, or  
XX fragment. The method is useful e.g. for identifying polynucleotides which  
XX singly or collectively encode intracellular immunoglobulin molecules, or  
XX which sensitize host cells to killing by an agent. The method may also be  
XX used in cardiovascular applications; for screening for diminished  
XX arrhythmia potential in cardiomyocytes and for enhanced contractile  
XX properties of cardiomyocytes and diminish heart failure potential; for  
XX identifying intracellular immunoglobulin molecules that will regulate  
XX intracellular and sarcolemmal calcium cycling in cardiomyocytes to  
XX prevent arrhythmias or that will diminish embolic phenomena in arteries

CC and arterioles leading to strokes and angina; in screening for decreases  
 CC in atherosclerosis-producing mechanisms to find intracellular  
 CC immunoglobulin molecules that regulate LDL and HDL metabolism; in skin  
 CC biology applications; and in regulating or inhibiting keloid formation.  
 CC AB22379 to AB22449 and AB25536 to AB25618 represent sequences used in  
 CC the exemplification of the present invention.

XX Sequence 32 AA;

Query Match 42.4%; Score 92; DB 24; Length 32;  
 Best Local Similarity 61.3%; Pred. No. 6.8e-07;

Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAALRAKHAHTSMVNFYRGKPV 31

Db 1 MLFNLRXXLNNAAFRHGHFMVNFRCGQPL 31

RESULT 10

AAB15704

ID AAB15704 standard; Peptide; 31 AA.

XX AC AAB15704;

DT 08-JAN-2001 (first entry)

DE Mitochondrial matrix localisation sequence.

KW Mitochondrial matrix localisation sequence; single-chain antibody;

KW stabilin; stabilising fusion peptide; vaccine; gene therapy;

KW protein degradation modulation; protein stability; Alzheimer's disease.

XX OS Unidentified.

XX FH Key Location/Qualifiers

FT Misc-difference 7 /note= "unspecified amino acid"

FT Misc-difference 31

FT /note= "unspecified amino acid"

XX WO200042185-A1.

XX PD 20-JUL-2000.

XX PF 11-JAN-2000; 2000WO-US00558.

XX PR 11-JAN-1999; 99US-0115505.

XX PA (MIND-) MINDSET BIOPHARMACEUTICALS USA INC.

XX PA (MCIN/) MCINNIS P A.

XX PI Chain DG;

XX DR WPI; 2000-476059/41.

XX Modulating degradability of protein or peptide useful for gene therapy  
 PT involving antibodies, comprises altering a gene at the N terminus to  
 PT render protein or peptide metabolically stable.

PS Disclosure; Page 50; 76pp; English.

XX The present sequence is a localisation sequence that can be used to  
 CC direct stabilised single-chain antibodies to the mitochondrial matrix.  
 CC The N-terminus of the single-chain antibody is linked to a stabilising  
 CC fusion peptide, referred to as a stabilon, which increases stability of  
 CC the antibody against proteolysis in vivo. Degradation of the antibody  
 CC may be modulated by linking the stabilon to the antibody through a  
 CC protease-sensitive linker region. The stabilon is removed upon induction  
 CC of expression of a specific restriction protease by means of an  
 CC inducible promoter, and this renders the antibody susceptible to  
 CC proteolysis by the N-end rule pathway. This method for regulating  
 CC protein stability allows removal of the antibody after it has bound to  
 CC its target antigen. Stabilised recombinant proteins may be used in gene

CC therapy for the treatment of disorders such as Alzheimer's disease.

XX SQ Sequence 31 AA;

Query Match 36.6%; Score 79.5; DB 21; Length 31;  
 Best Local Similarity 61.3%; Pred. No. 5.8e-05;

Matches 19; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

QY 1 MLSNRLILNKAALRAKHAHTSMVNFYRGKPV 31

Db 1 MLFNLRXXLNNAAFRHGHFMVNFRCGQPL 30

RESULT 11

AAB22835

ID AAB22835 standard; peptide; 31 AA.

XX AC AAB22835;

DT 10-JAN-2001 (first entry)

DE Mitochondrial matrix localisation signal.

KW Controlled release delivery system; drug targeting;

KW drug-specific antibody; intracellular half-life; gene therapy;

KW diabetes; autoimmune disease; inflammatory disease; infectious disease;

KW cancer; side effect; subcellular localisation sequence.

XX OS Unidentified.

XX FH Key Location/Qualifiers

FT Misc-difference 7 /label= unknown

FT Misc-difference 31

FT /label= unknown

XX WO200050089-A2.

XX PD 31-AUG-2000.

XX PF 25-FEB-2000; 2000WO-US04749.

XX PR 26-FEB-1999; 99US-0122103.

XX PA (MIND-) MINDSET BIOPHARMACEUTICALS USA INC.

XX PI Chain DG;

XX DR WPI; 2000-572044/53.

XX Delivering a drug, to a patient suffering from cancer or diabetes, at a  
 PT predetermined site, comprises altering a target cell to express a drug  
 PT specific antibody by gene therapy and administering the drug.

PS Disclosure; Page 44; 69pp; English.

XX The invention relates to a controlled release drug delivery system.  
 CC Target cells in a patient are transformed with an expression construct  
 CC encoding an antibody (particularly a single chain antibody) to the drug  
 CC to be delivered. The antibody contains a modulator of intracellular  
 CC half-life: this can either be a stabilising or destabilising residue  
 CC located in the N-terminus after the initial methionine, or a peptide  
 CC ("stabilon") containing a stabilising residue linked to the antibody  
 CC N-terminus via a protease cleavage site. On administration of the drug,  
 CC the antibody binds the drug, localising it at the target cells but  
 CC maintaining it in an inactive state. As the antibody is degraded (the  
 CC timescale for which is dependent upon the N-terminal or stabilon amino  
 CC acids), the drug is released at its site of action where it can exert  
 CC its effects. The antibody encoded by the expression construct (and  
 CC therefore the therapeutic agent) can be targeted to particular  
 CC subcellular locations (e.g., the nucleus) by including the appropriate  
 CC cellular localisation signals. The novel method may be used to deliver  
 CC therapeutic agents to patients with a variety of conditions such as

diabetes, autoimmune diseases, inflammatory diseases, infectious diseases and especially cancer. The delivery method of the invention causes a drug to become almost immediately localised at its site of action in an inactive form where it accumulates. Once released by antibody degradation, the drug is at an effective concentration only at the target site, with very little free drug being available in the rest of the body. The system of the invention therefore reduces the side effects caused by therapeutic agents, and also provides economic benefits as a smaller amount can be administered to the patient. Sequences AAB22833 - AAB22837 and AAB22839-822857 represent subcellular localisation sequences which can be incorporated into a drug-specific antibody used in the method of the invention.



```

ABE59259
ID  ABB59259 standard; Protein; 755 AA.
XX
AC  ABB59259;
XX
DT  26-MAR-2002 (first entry)
XX
DE  Drosophila melanogaster polypeptide SEQ ID NO 4569.
XX
KW  Drosophila; developmental biology; cell signalling; insecticide;
KW  Pharmaceutical.
XX
OS  Drosophila melanogaster.
XX
PN  WC2000171042-A2.
XX
PD  27-SEP-2001.
XX
PF  23-MAR-2001; 2001WO-US09231.
XX
PR  23-MAR-2000; 2000US-191637P.
PR  11-JUL-2000; 2000US-0614150.
XX
PA  (PEKE ) PE CORP NY.
XX
PI  Venter JC, Adams M, Li PWD, Myers EW;
XX
DR  WPI; 2001-656860/75.
DR  N-PSDB; ABL03362.
XX
XX  New isolated nucleic acid detection reagent for detecting 1000 or more
PT  genes from Drosophila and for elucidating cell signalling and cell-cell
PT  interactions.
XX
PS  Disclosure; SEQ ID NO 4569; 21pp + Sequence Listing; English.
XX
CC  The invention relates to an isolated nucleic acid detection reagent
CC  capable of detecting 1000 or more genes from Drosophila. The invention is
CC  useful in developmental biology and in elucidating cell signalling and
CC  cell-cell interactions in higher eukaryotes for the development of
CC  insecticides, therapeutics and pharmaceutical drugs. The invention
CC  discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC  sequences (ABL01840-ABL16175) and the encoded proteins
CC  (ABE57737-ABE72072).
CC  The sequence data for this patent did not form part of the printed
CC  specification, but was obtained in electronic format directly from WIPO
CC  at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ  Sequence 755 AA;

Query Match      24.7%; Score 53.5; DB 22; Length 755;
Best Local Similarity 42.5%; Pred. No. 33;
Matches 17; Conservative 5; Mismatches 15; Indels 3; Gaps 2;

QY  2 LSNRLTLNKAALRKRAHTSMVNFYRFGKPVQSQVOLKPRD 41
   ||::||| | ||||| : : || : : ||||
Db  11 LSDLRRELE--VLRKAHFELDHLFYGTG-QPEAEAKPRD 47

RESULT 15
AAY04933
ID  AAY04933 standard; Protein; 165 AA.
XX
AC  AAY04933;
XX
DT  06-JUL-1999 (first entry)
XX
DE  Mycobacterium species protein sequence 388.
XX
KW  Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW  hybridisation; detection; vaccine; immunisation; infection.
XX
OS  Mycobacterium sp.

XX  WO9909186-A2.
XX
XX  25-FEB-1999.
XX
XX  14-AUG-1998; 98WO-FR01813.
XX
XX  11-SEP-1997; 97FR-0011325.
XX  14-AUG-1997; 97FR-0010404.
XX
XX  (INSP ) INST PASTEUR.
XX
XX  Gicquel B, Lim EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y;
PI  Guigueno A;
XX
XX  WPI; 1999-181045/15.
XX  N-PSDB; AAX34186.
XX
XX  Mycobacterial DNA vectors containing reporter constructs - for
PT  identifying coding or promoter sequences involved in
PT  infection-associated protein expression
XX
PS  Claim 32; Fig 38B; 309pp; French.
XX
XX  Sequences AAY04742-Y05000 and AAY07201-Y07204 represent secreted
CC  proteins from various Mycobacterium species microorganisms. The
CC  encoding nucleotide sequences can be used as primers and probes for
CC  methods for detecting and identifying mycobacteria, especially belonging
CC  to the M. tuberculosis complex. The encoded proteins can be used in
CC  vaccines for immunisation against a bacterial or viral infection.
XX
SQ  Sequence 165 AA;

Query Match      24.0%; Score 52; DB 20; Length 165;
Best Local Similarity 37.1%; Pred. No. 8.9;
Matches 13; Conservative 8; Mismatches 12; Indels 2; Gaps 1;

QY  6 RILNKAALRKRAHTSMVNFYRFGKPVQSQVOLKPR 40
   ||:| | |||: | | | : : |||
Db  119 RVILRCATRKANQSRARTLRPL--RIALRPR 151

Search completed: July 24, 2003, 22:04:08
Job time : 38.881 secs

```